

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:15:08 ; Search time 163 seconds

(without alignments)
1657.775 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSILVLCNV.....VDLKKDCRRRLRDPFPCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	383	2	Q9GUM2 CAEBL
2	1848.5	91.0	384	2	Q60PPO CAEBR
3	775.5	38.2	421	2	Q6J4T9 TRINI
4	743	36.6	260	2	Q7Q2R7 ANOGA
5	736.5	36.3	498	2	Q6NU80 XENIA
6	730	35.9	498	2	Q8AVL9 XENIA
7	729	35.9	336	2	Q6DDG6 XENIA
8	729	35.9	354	2	Q6DIT3 XENTR
9	725	35.7	395	1	B4GT3 MOUSE
10	720	35.5	354	2	Q5EPY8 XENIA
11	720	35.5	396	1	B4GT3 BOVIN
12	717	35.3	393	1	B4GT3 HUMAN
13	717	35.3	393	1	B4GT3 PONPY
14	717	35.3	395	1	B4GT3 CRIGR
15	716	35.3	395	1	B4GT3 RAT
16	709	34.9	403	2	Q9XZ05 DROME
17	696	34.3	360	2	Q4T5R7 TENG
18	694	34.2	395	2	Q71VA3 HUMAN
19	680.5	33.5	362	2	Q92074 CHICK
20	679.5	33.5	341	2	Q4SFL2 TENG
21	671.5	33.1	393	2	Q99PCL CRIGR
22	670	33.0	373	2	Q92073 CHICK
23	667.5	32.9	399	1	B4GT1 MOUSE
24	666.5	32.8	413	2	Q9NDP2 CLOIN
25	666	32.8	397	1	B4GT1 HUMAN
26	660	32.5	372	2	Q4V9L9 HUMAN
27	657	32.3	372	1	B4GT2 HUMAN
28	655.5	32.3	353	2	Q5UJQ4 BRARE
29	651.5	32.1	402	1	B4GT1 BOVIN
30	651	32.1	344	1	B4GT4 CRIGR
31	648.5	31.9	369	1	B4GT2 CRIGR

32	646.5	31.8	369	1	B4GT2 MOUSE	Q92Y2 m beta-1,4-
33	642.5	31.6	344	1	B4GT4 HUMAN	Q60513 h beta-1,4-
34	635	31.3	344	1	B4GT4 RAT	Q661h1 r beta-1,4-
35	633	31.2	344	1	B4GT4 MOUSE	Q9J104 m beta-1,4-
36	632	31.1	430	2	Q4S57 TENG	Q4S57 tetradon n
37	632	31.1	557	2	Q4P16 TENG	Q4P16 tetradon n
38	629	31.0	310	2	Q64U7 PIG	Q64u7 sus scrofa
39	621.5	30.6	347	2	Q6GLP1 XENIA	Q6glp1 xenopus lae
40	616	30.3	319	2	Q6BKO BRARE	Q6Bko brachydanio
41	596.5	29.4	382	2	Q4RLE6 TENG	Q4rle6 tetradon n
42	593	29.2	368	2	Q5F35 CHICK	Q5f35 gallus gal1
43	584.5	28.8	388	1	B4GT5 MOUSE	Q9jmk0 m beta-1,4-
44	584.5	28.8	388	1	Q80W6 CRIGR	Q80w6 cricetus
45	584	28.8	382	2	Q4SDX3 TENG	Q4sdx3 tetradon n

ALIGNMENTS

RESULT 1
ID Q9GUM2 CAEBL PRELIMINARY; PRT; 383 AA.
AC Q9GUM2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE Bt (Bacillus thuringiensis) toxin resistant protein 4
DE (UDPGalNAc:GlcNAc[beta]-R (beta)1,4-N-acetylgalactosaminyl[transferase]
DE (BRE-4).
GN Name=bre-4; ORFNames=Y73E7A.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodietinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN NM_011111
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
[2]
RN NM_011111
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22218093; PubMed=12167666; DOI=10.1074/jbc.M206112200;
RA Kawai S.S., Van Die I., Cummings R.D.,
RT "Molecular Cloning and Enzymatic Characterization of a UDP-
RT GalNAc:GlcNAc[beta]-R beta 1,4-N-Acetylgalactosaminyltransferase from
RT Caenorhabditis elegans.";
RL J. Biol. Chem. 277:34924-34932(2002).
[3]
RN NM_011111
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22970005; PubMed=12944392; DOI=10.1074/jbc.M308142200;
RA Griffiths J.S., Huffman D.L., Whitacre J.L., Barrows B.D.,
RA Marroquin L.D., Muller R., Brown J.R., Hennes T., Esko J.D.,
RA Arojan R.V.,
RT "Resistance to a bacterial toxin is mediated by removal of a conserved
RT glycosylation pathway required for toxin-host interactions.";
RL J. Biol. Chem. 278:45594-45602(2003).
[4]
RN NM_011111
RP NUCLEOTIDE SEQUENCE.
RA Griffiths J.S., Huffman D.L., Arojan R.V.,
RL Submitted (JAN-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AC025727; AAG3384.1; -; Genomic DNA.
DR EMBL; AY130767; AAM95168.1; -; mRNA.
DR HSSP; P08037; 100R.
DR Ensemble; Y73E7A.7; Caenorhabditis elegans.
DR WormBase; WBGen0000269; bre-4.
DR WormPeP; Y73E7A.7; CE26412.
GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR003859; Galactosyl_T_2.

DR InterPro: IPR001998; Xylose_isom.
DR Pfam: PR02709; Galactosyl_T_2; 1.
DR PROSITE: PS00173; XYLOSE_ISOMERASE_2; UNKNOWN_1.
KW Complete proteome; transferase.
SQ SEQUENCE 383 AA; 43914 MW; ADEDF44275B2CE81 CRC64;
Query Match 100.0%; Score 2031; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 3,4e-165;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAFRLHVARLKSLLVLCALLLVHAMIKYKIPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
DB 1 MAFRLHVARLKSLLVLCALLLVHAMIKYKIPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
QY 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFCNQTPPHLVGPIRVFLD 120
DB 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFCNQTPPHLVGPIRVFLD 120
QY 121 PPFKTEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDREAHRLIMLHNLHSLAKQQLD 180
DB 121 PPFKTEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDREAHRLIMLHNLHSLAKQQLD 180
QY 181 VAIPIVEQVANQTFNRGKLMNVGYDVASRLYPWCFFIHDVDLLPEDDRNLVYTCPIQPR 240
DB 181 VAIPIVEQVANQTFNRGKLMNVGYDVASRLYPWCFFIHDVDLLPEDDRNLVYTCPIQPR 240
QY 241 MSVAIDKFNKYLPSAIFGGISALTQKHLKINFSNDFPMWGGEEDDLARTSMAGLV 300
DB 241 MSVAIDKFNKYLPSAIFGGISALTQKHLKINFSNDFPMWGGEEDDLARTSMAGLV 300
QY 301 SRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGOTKRWTRDGLSNLKYKLVNLEKPLY 360
DB 301 SRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGOTKRWTRDGLSNLKYKLVNLEKPLY 360
QY 361 RAVVDLLEKDCRRELRRDPFTCF 383
DB 361 RAVVDLLEKDCRRELRRDPFTCF 383
RESULT 2
Q60PPO CAEBR PRELIMINARY; PRT; 384 AA.
AC Q60PPO;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Hypochemical protein CBG22165.
GN Name=CBG22165;
OS Caenorhabditis briggsae.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000129; CAE74432.1; -; Genomic DNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR001998; Galactosyl_T_2.
DR InterPro: IPR001998; Xylose_isom.
DR Pfam; PF02709; Galactosyl_T_2; 1.
DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; UNKNOWN_1.
KW Hypochemical protein.
SQ SEQUENCE 384 AA; 43811 MW; 4P9BEC583A634B7C CRC64;
Query Match 91.0%; Score 1848.5; DB 2; Length 384;
Best Local Similarity 90.1%; Pred. No. 1.5e-149;
Matches 346; Conservative 16; Mismatches 21; Indels 1; Gaps 1;

QY 1 MAFRLHVARLKSLLVLCALLLVHAMIKYKIPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
DB 1 MAFRLHVARLKSLLVLCALLLVHAMIKYKIPSLYENLTIGSSSTLIADVDAMEAVLGNTA 60
QY 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFCNQTPPHLVGPIRVFLD 119
DB 61 STSDLLDTWNSTFSPISSEVNOTSFMEDIRPLIFPNQTLQFCNQTPPHLVGPIRVFLD 120
QY 120 EPDFKTEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDREAHRLIMLHNLHSLAKQQL 179
DB 120 EPDFKTEKTYPDTHAGGHGMPKDCVARRHVAIIVPYRDREAHRLIMLHNLHSLAKQQL 180
QY 181 DVAIPIVEQVANQTFNRGKLMNVGYDVASRLYPWCFFIHDVDLLPEDDRNLVYTCPIQPR 240
DB 181 DVAIPIVEQVANQTFNRGKLMNVGYDVASRLYPWCFFIHDVDLLPEDDRNLVYTCPIQPR 240
QY 240 HNSVAIDKFNKYLPSAIFGGISALTQKHLKINFSNDFPMWGGEEDDLARTSMAGLV 299
DB 240 HNSVAIDKFNKYLPSAIFGGISALTQKHLKINFSNDFPMWGGEEDDLARTSMAGLV 300
QY 300 VSRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGOTKRWTRDGLSNLKYKLVNLEKPLY 359
DB 301 VSRPPTQIARYKMIKHSSTEATNPVNCRCRYKIMGOTKRWTRDGLSNLKYKLVNLEKPLY 360
QY 360 TRAVVDLLEKDCRRELRRDPFTCF 383
DB 361 TRAVVDLLEKDCRRELRRDPFTCF 384
RESULT 3
Q6J4T9 TRINI PRELIMINARY; PRT; 421 AA.
AC Q6J4T9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Beta 1,4-N-acetylgalactosaminyltransferase.
OS Trichoplusia ni (Cabbage looper).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Dityraria; Noctuoidea;
OC Noctuidae; Plusiinae; Trichoplusia.
OX NCBI_TaxId=7111;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15173167; DOI=10.1074/jbc.M404925200;
RA Vadate N., Jarvis D.L.,
RT "Molecular Cloning and Functional Characterization of a Lepidopteran
RT Insect (beta)4-N-acetylgalactosaminyltransferase with Broad Substrate
RT Specificity, a Functional Role in Glycoprotein Biosynthesis, and a
RT Potential Functional Role in Glycolipid Biosynthesis.",
RL J. Biol. Chem. 279:33501-33518(2004).
DR EMBL; AY601103; AAT11926.1; -; mRNA.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro: IPR003859; Galactosyl_T_2.
DR InterPro: IPR001998; Xylose_isom.
DR Pfam; PR02709; Galactosyl_T_2; 1.
DR PROSITE; PS00173; XYLOSE_ISOMERASE_2; UNKNOWN_1.
KW Transferase.
SQ SEQUENCE 421 AA; 48302 MW; EFF3CC1168BBD0AB CRC64;
Query Match 38.2%; Score 775.5; DB 2; Length 421;
Best Local Similarity 43.5%; Pred. No. 1.3e-57;
Matches 157; Conservative 63; Mismatches 116; Indels 25; Gaps 6;
QY 30 KIPSLY-----ENLTIGSSTLIADVDAMEAVLGNTASTSDLLDTWNSTFSPISSEV 80
DB 64 KIPSNYIEDSESEYKINISLSNHTTRASVYHPPSSITETASKLDKMTTIDGAFAMISPT 123
QY 81 NQ--TSPMEDIRPLIPDN-----QTLQFCNQTPPHLVGPIRVFLDEPDKTEKTY 130
DB 124 PLITIKMDSIKSYVTTEGKVKAQAVVTLPLCDSPMDL-GPITLTKTELEWEVKCF 182

QY 131 PHTAGHGMPKDCVARRHVAIIVYRDEAHRLIMLHSLAKQQLDVAIFIVEQVA 190
DB 133 FEVEGGRSPNCTARRHVAIIVYRDBQQLALFLNMHPFLKQQLVEYGFIVEQGG 242
QY 191 NOTFRGKLMNVGYVASRLYF--WQCFIFHDVLLPEDRRLYTCPIOPRMSVAIDKF 248
DB 243 NMDFRKALMNVGFESQCLVAEGWQCFVHDIIDLPLDTRNLVSCPRPRMSASIDEL 302
QY 249 NYKLFPYSAIFGSIALTQDKLKKINGFSDPFMGWGEDDDLLATRTSMGLKYSRYPTOLA 308
DB 303 HFKLPYEDIFGVSAMTLEQFTRVNGFSNKGWGEDDDMSYRLKKNYHIAKRMSTA 362
QY 309 RYKMKHSTKTEATNPVNCGRYKIMGOTKRRWTRDGLSNLYKLVNLEKPLYTRAVDLE 368
DB 363 RYAMLDKHSTPNPK---RYQLSQTSKTFQDGLSTLELYVQVYHLIHLINIDE 419
QY 369 K 369
DB 420 R 420

RESULT 4

Q7Q2R7_ANOGA PRELIMINARY; PRT; 260 AA.
ID Q7Q2R7;
AC Q7Q2R7;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGF0000010877 (fragment).
GN ORFNames=ENSANG0000008388;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anophelinae.
NCBI_TaxID=180454;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG "Anopheles gambiae re-annotation";
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
CC -i- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AAB01008968; EAA13217.2; -; Genomic_DNA.
DR InterPro: IPR003959; Galactosyl_T_2.
DR InterPro: IPR001998; Xylose_Isom.
DR Pfam: PF02709; Galactosyl_T_2; 1.
DR PROSITE: PS00173; XYLOSE_ISOMERASE_2; UNKNOWN_1.
FT NON TER 1 260
FT 260
SQ SEQUENCE 260 AA, 29720 MW; 24B8F310B4781932 CRC64;

Query Match 36.6%; Score 743; DB 2; Length 260;
Best Local Similarity 54.3%; Pred. No. 4e-55;
Matches 139; Conservative 36; Mismatches 77; Indels 4; Gaps 2;

QY 112 GRIFFLDEPDKTEKLYPD--THAGHMPKDCVARRHVAIIVYRDEAHRLIMLHSL 170
DB 1 GRIYDVAFESISAVBSRPADKLQPGQVVPDCTARRHVAIIVYRDEKHLPIFLKNI 60
QY 171 HSLAKQQLDVAIFIVEQVANTFRNGKLMNVGYVASRLYF--WQCFIFHDVLLPEDDRN 230
DB 61 HALLMKQQLVEYGFIVEQVASSFNRAALMNTGFVEAMQKMECWFHDDLLMDDDN 120
QY 231 LYTGPDPHMSVAIDKFNYKLYSAIFGSIALTQDKLKKINGFSDPFMGWGEDDDLA 290
DB 121 LYTCDPDPHMSVAIDVFGLKLYSTIFGVSAMTEKQPRMNVNGFSNKGWGEDDDMS 180

QY 291 TRTSMAGLVERYPQIARYKMKHSTKTEATNPVNCGRYKIMGOTKRRWTRDGLSNLYKYL 350
DB 181 NRKLVGHRIARYPNINARIYMLSHKKEKANPK---RREKLVNAGKRDSGLSLHQL 237
QY 351 VNLEKPLYTRAVDLE 366
DB 238 VNLRKPLYTWIHADI 253

RESULT 5

Q6NU80_XENLA PRELIMINARY; PRT; 498 AA.
ID Q6NU80;
AC Q6NU80;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE MGC81163 protein.
GN Name=MGC81163;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
NCBI_TaxID=8355;
[1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Ovary;
RC MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Klausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausberg R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Martina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.I., Skalski U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RA Klein S., Strausberg R.;
RL Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC068719; AAH68719.1; -; mRNA.
DR GO: GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro: IPR003859; Galactosyl_T_2.
DR Pfam: PF02709; Galactosyl_T_2; 1.
SQ SEQUENCE 498 AA; 56855 MW; 5A322F40FEDOC01F CRC64;

Query Match 36.3%; Score 736.5; DB 2; Length 498;
Best Local Similarity 44.6%; Pred. No. 3.5e-54;
Matches 149; Conservative 54; Mismatches 114; Indels 17; Gaps 6;

DT 13-SBP-2005 (Rel. 48, Last sequence update)
 DT 13-SBP-2005 (Rel. 48, Last annotation update)
 DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.1) (Beta-1,4-GalTase 3)
 DE (Beta4Gal-1.3) (b4Gal-1.3) (UDP-galactose:beta-N-acetylglucosamine beta-
 1,4-galactosyltransferase 3) (UDP-Gal:beta-GlcNAc beta-1,4-
 galactosyltransferase 3) [includes: N-acetylglucosamine synthase
 (EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide
 beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-
 acetylglucosaminyl glycolipid beta-1,4-galactosyltransferase
 (EC 2.4.1.1-1)].
 DE Name=B4GalT3;
 GN Name=B4GalT3;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RX NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RX MEDLINE=98259838; PubMed=9597550; DOI=10.1093/glycob/8.5.517;
 RA Lo N.-W., Shaper J.H., Pevsner J., Shaper N.L.;
 RT "The expanding beta 4-galactosyltransferase gene family: messages from
 the databanks.";
 RL Glycobiology 8:517-526(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.F., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant P.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulvaney S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fehy J., Hulton E., Kettelman M., Madan A.C., Shvachenko Y., Bouffard G.G.,
 RA Whiting M., Madan A., Young A.C., Shvachenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Responsible for the synthesis of complex-type N-linked
 oligosaccharides in many glycoproteins as well as the carbohydrate
 moieties of glycolipids (By similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-
 glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-
 beta-D-glucosaminylglycopeptide.
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +
 N-acetylglucosamine.
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of
 Golgi stack (By similarity).
 CC -1- SIMILARITY: Belongs to the glycosyltransferase 7 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: AF142671; AAF22221.1; -; mRNA.
 CC EMBL: BC013619; AAH13619.1; -; mRNA.
 DR HSSP: P08037.1RGX.
 DR Ensembl: ENSMUSG0000052423; Mus musculus.
 DR MGI: MGI:1928767; B4galT3.
 GO: GO:0016021; C:integral to membrane; TAS.

DR InterPro: IPR003859; Galactosyl T 2.
 DR PANTHER: PTHR19300; Galactosyl T 2; 1.
 DR Pfam: PF02709; Galactosyl T 2; 1.
 KW Glycoprotein; Glycosyltransferase; Golgi stack; Manganese;
 KW Metal-binding; Multigene family; Signal-anchor; Transferase;
 KW Transmembrane.
 FT TOPO_DOM 1 10 Cytoplasmic (Potential).
 FT TRANSMEM 11 31 Signal-anchor for type II membrane
 FT FT protein (Potential).
 FT TOPO_DOM 32 395 Luminal (Potential).
 FT METAL 199 199 Manganese (By similarity).
 FT METAL 293 293 Manganese (By similarity).
 FT CARBOHYD 57 57 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 168 168 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 339 339 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 387 387 N-linked (GlcNAc...) (Potential).
 FT DISULFID 79 121 By similarity.
 FT DISULFID 192 211 By similarity.
 FT CONFID 353 353 S -> A (in Ref. 1).
 SQ SEQUENCE 395 AA; 44084 MW; 16401B03EB4B21 CRC64;
 Query Match 35.7%; Score 725; DB 1; Length 395;
 Best Local Similarity 48.9%; Pred. No. 2.5e-53;
 Matches 134; Conservative 44; Mismatches 94; Indels 2; Gaps 2;
 QY 95 PDNQTQPCNQTPPHLVGPPIRV-FLDEPDKLTKIYPTTHAGHGMPKDCVARRHVAII 153
 DB 71 PPAQALPYCPBESFPLVGVSFSFVPSLAELVEGNPRVSGGRPRAPGCEPRSTAI 130
 QY 154 VPYRDEAHLRIHLNHLAKQQIDVAIFVEQVANOFTFRGKLMNVGYDVASFLYPM 213
 DB 133 VPRAREHHLRIILYHLHPFLORQQLAAGIYVHQNGTFRNRAKLNVGEALDEEM 190
 QY 214 QCFIFPDVLLPDDSNLYTC-PIQPRHMSVMIIDKENVLPISALFGISALTKDKLKKI 272
 DB 191 DCLFLHDVLLPDDSNLYTC-PIQPRHMSVMIIDKENVLPISALFGISALTKDKLKKI 250
 QY 273 NGFSNDFMGCGEDDILATRTSNAGLKVGRVPTQIARVYKMIKSTATNPVVKCRKYNG 332
 DB 251 NGFPRNFMGCGEDDILATRTSNAGLKVGRVPTQIARVYKMIKSTATNPVVKCRKYNG 310
 QY 333 QTRKRWTRDGLSNLYKLVNLKIPYTRAVVDL 366
 DB 311 RTQNSWTQDGMNSLTYRLRLARELGPITYNTADI 344
 RESULT 10
 Q5PPY8_XENILA PRELIMINARY; PRT; 354 AA.
 AC Q5PPY8;
 DT 01-FEB-2005 (T-EMBLrel. 29, Created)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
 DE LOC436041 protein.
 GN Name=LOC436041;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodidae; Xenopus; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT initiative.";
 RL Dev. Dyn. 225:384-391(2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haich F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Wang C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Planchy S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettlemen M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Testis;
 RA Klein S., Gerhard D.S.;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC087436; AAH87436.1; -; mRNA.
 DR GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
 DR CO; GO:0005975; P:carbohydrate metabolism; IEA.
 DR InterPro; IPR003859; Galactosyl_T_2.
 DR Pfam; PF02709; Galactosyl_T_2; 1.
 SQ SEQUENCE 354 AA; 40970 MW; 016170773E16B208 CRC64;

Query Match 35.5%; Score 720; DB 2; Length 354;
 Best Local Similarity 40.3%; Pred. No. 5,7e-53;
 Matches 143; Conservative 61; Mismatches 117; Indels 34; Gaps 5;
 14 LTVLCVLLVHAMIYK--IPSLYENLTIGSSSTLADVDAMEVLGNTASTSDLLDPTN 71
 12 LLLLCFSQLLFLILYRGAAGVQGLRGGSD-----PPMD 47
 72 STFSPISEVNGTSPMEDIRPILFPNDQLQFCNQPPHVGPIRV-FLDEPDKLEKTY 130
 48 YEKTHDVYTNLSVFVN-----PENVNQYCPKAPILVGPLVSFGPPPTLKRVQKN 101
 131 PPTHAGCHMPKDCVARRHVAIIVPRDBEAHRLMNLHLSLAKOQDVAIFIVEQVA 190
 102 RYVKPGGYSPRRHCFGRYRTAVIIPRRNEAHLRTLLYVLPFLQROQLVAFIVHQAG 161
 191 NOTFNRGLMNVGYDVASRLYPMQCFIPHDVDLLPEDDRNLTYTCPIQ-PRHMSVAIDKEN 249
 162 NSTFNRKALNLGVRRALKLDEMDCILHLDVLPENDTNLYICDEEYKHLASAMDKFH 221
 250 YKLPSAIFGIGISALTQDHLKKINGFNSDFWGWGEGDDDLATRTSMAGLKVSRYPQIAR 309
 222 YSLPYTWYGGVSALPPDQYMRINGEPNPGWGGEGDDDIAMRIRLAGMSITRTPLSLGR 281
 310 YNMIGSTATATPVKCKRYKIMGQTKRKTTRDGLSLUKLVNLEKPIYTAAV 364
 282 YKMISHNRDSGNENSKRYDOLGNTRTWRDEGMNSLDKFLISRTAPLTYNITV 336

RESULT 11
 BACT3_BOVIN STANDARD; PRT; 396 AA.
 AC Q5EA87; Q5E9K4;
 DT 13-SBP-2005 (Rel. 48, Created)
 DT 13-SBP-2005 (Rel. 48, Last sequence update)
 DT 13-SBP-2005 (Rel. 48, Last annotation update)
 DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.1.-) (Beta-1,4-GalTase 3)
 DE (Beta4Gal-1T3) (b4Gal-T3) (UDP-galactose:beta-N-acetylglucosamine beta-
 1,4-galactosyltransferase 3) (UDP-gal:beta-GlcNAc beta-1,4-
 galactosyltransferase 3) [Includes: N-acetylglucosamine synthase
 (EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide

DE beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-
 DE acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase
 DE (EC 2.4.1.-)).
 GN Name=B4GALT3;
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
 OC Pecora; Bovidae; Bovinae; Bos.
 OX NCBI_TaxId=9913;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
 RA Harhay G.P., Sonstegard T.S., Clawson M.L., Heaton M.P., Keele J.W.,
 RA Snelling W.M., Weidman R.T., Smith T.P.L.;
 RT "Sequencing and analysis of Bos taurus full-length insert cDNA
 clones.";
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Responsible for the synthesis of complex-type N-linked
 CC oligosaccharides in many glycoproteins as well as the carbohydrate
 CC moieties of glycolipids (By similarity).
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-
 CC glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-
 CC beta-D-glucosaminylglycopeptide.
 CC -1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +
 CC N-acetylglucosamine.
 CC -1- COFACTOR: Manganese (By similarity).
 CC -1- PATHWAY: Glycosylation.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of
 CC Golgi stack (By similarity).
 CC -1- SIMILARITY: Belongs to the glycosyltransferase 7 family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC EMBL; BT020682; AAX08699.1; -; mRNA.
 CC EMBL; BT020686; AAX08885.1; -; mRNA.
 CC EMBL; BT020916; AAX08933.1; -; mRNA.
 DR InterPro; IPR003859; Galactosyl_T_2.
 DR PANTHER; PTHR19300; Galactosyl_T_2; 1.
 DR Pfam; PF02709; Galactosyl_T_2; 1.
 KW Glycoprotein; Glycosyltransferase; Golgi stack; Manganese;
 KW Metal-binding; Multigene family; Signal-anchor; Transferase;
 KW Transmembrane.
 FT TOPO_DOM 1 10 Cytoplasmic (Potential).
 FT TRANSMEM 1 31 Signal-anchor for type II membrane
 FT TOPO_DOM 1 31 protein (Potential).
 FT METAL 200 200 Manganese (By similarity).
 FT METAL 294 294 Manganese (By similarity).
 FT CARBOHYD 57 57 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 169 169 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 340 340 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 388 388 N-linked (GlcNAc...) (Potential).
 FT DISULFID 80 122 By similarity.
 FT DISULFID 193 212 By similarity.
 FT CONFLICT 82 82 K -> E (in Ref. 1; AAX08699).
 SQ SEQUENCE 396 AA; 44371 MW; 8C2611ED7C26250D CRC64;

Query Match 35.5%; Score 720; DB 1; Length 396;
 Best Local Similarity 48.6%; Pred. No. 6,7e-53;
 Matches 136; Conservative 44; Mismatches 98; Indels 2; Gaps 2;
 89 IRPILFPNDQLQFCNQPPHVGPIRV-FLDEPDKLEKTYPTTHAGCHMPKDCVARR 147
 66 VAPGGLPAPGGLPYCPKRSPLLVGISVSFSPVSLAIVERNPRVPGGRYPARCEPR 125
 148 HRVAIIVPRDBEAHRLMNLHLSLAKOQDVAIFIVEQVANTFNRGLMNVGYDVA 207
 126 STTAIVPRAREHRLRLVLPFLQROQLVAFIYVHQAGTNPFRKALNLGVARRA 185
 208 SRLYPMQCFIPHDVDLLPEDDRNLTYTC-PIOPRHMSVAIDKENYKLPYSAIFGIGISALTQK 266

DB 186 LRDEBDC.FLHDVDDLPERDNLVYCDPGRPHVAVAMNKFGXSLPYQFGVGSALRP 245
 QY 267 DHAKTNGSNDPWGCGEDDLATRTSMGLKVSRYPOIAAYKMIKSTKNTNVNKC 326
 DB 246 DQYLNKGFPNEXWGWGDEDDDLATRVRLGKMISSPPTSVGHYKWKHGXGNEENPH 305
 QY 327 RYKMGOTKKRRWTRDGLSNLKYLVNLEPLVTRAVVDL 366
 DB 306 RPDVLVTRTONSWTQDGMNSLTLYQLLSRELGLPTNTITADI 345
 RESULT 12
 BACT3 HUMAN STANDARD: PRT; 393 AA.
 AC 060512; 060910; Q9B8P2; Q9H8T2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Beta-1,4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-GalTase 3)
 DE Beta4Gal-T3 (b4Gal-T3) (UDP-galactose:beta-N-acetylglucosamine beta-1,4-galactosyltransferase 3) (UDP-Gal:beta-GlcNAc beta-1,4-galactosyltransferase 3) (Includes: N-acetylglucosamine synthase
 DE beta-1,4-galactosyltransferase) (EC 2.4.1.38); Beta-N-acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase
 DE (EC 2.4.1.-).
 GN Name=B4GALT3;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RX MEDLINE=98070356; PubMed=9405390; DOI=10.1074/jbc.272.51.31979;
 RA Almeida R., Amado M., David L., Levery S.B., Holmes E.H., Merx G.,
 RA van Kessel A.G., Rygaard E., Hassan H., Bennett E., Clausen H.;
 RT "A family of human beta4-galactosyltransferases. Cloning and
 RT expression of two novel UDP-galactose:beta-n-acetylglucosamine beta1,
 RT 4-galactosyltransferases, beta4Gal-T2 and beta4Gal-T3.";
 RL J. Biol. Chem. 272:31979-31991 (1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1).
 RX MEDLINE=98259838; PubMed=9597550; DOI=10.1093/glycob/8.5.517;
 RA Lo N.-W., Shaper J.H., Peyssner J., Shaper N.L.;
 RT "The expanding beta 4-galactosyltransferase gene family: messages from
 RT the databanks.";
 RL Glycobiology 8:517-526 (1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), AND BIOPHYSICOCHEMICAL
 RP PROPERTIES.
 RX MEDLINE=21472272; PubMed=11588157; DOI=10.1093/glycob/11.10.813;
 RA Guo S., Sato T., Shirane K., Furukawa K.;
 RT "Galactosylation of N-linked oligosaccharides by human beta-1,4-
 RT galactosyltransferases I, II, III, IV, V, and VI expressed in Sf-9
 RT cells.";
 RL Glycobiology 11:813-820 (2001).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RG Human chromosome 1 international sequencing consortium;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
 RC TISSUE=Ovary;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekine N., Ohtsue M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,

RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
 RA Yamazaki M., Ninomiya K., Ishibaishi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimeta M., Watanabe S., Yoshida S., Chiba Y.,
 RA Ieshida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houta T.,
 RA Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O.,
 RA Nomura Y., Togiyama S., Komai F., Hara R., Takuchi K., Arita M.,
 RA Imose N., Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terasima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Taniguchi A., Fujita Y.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtsuri R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Moringa M., Sasaki M.,
 RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isozaki T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45 (2004).
 RN [6]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).
 RC TISSUE=Brain, Eye, and Lymph;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins S.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udsell T.B., Tomshinsky S., Carninci P., Prange C.,
 RA Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mallya S.J.,
 RA Bosak S.A., McWain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
 RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Holton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [7]
 RP REVIEW.
 RX MEDLINE=20047730; PubMed=10580128; DOI=10.1016/S0304-4165(99)00168-3;
 RA Amado M., Almeida R., Schwientek T., Clausen H.;
 RT "Identification and characterization of large galactosyltransferase
 RT gene families: galactosyltransferases for all functions.";
 RL Biochim. Biophys. Acta 1473:35-53 (1999).
 CC - FUNCTION: Responsible for the synthesis of complex-type N-linked
 CC oligosaccharides in many glycoproteins as well as the carbohydrate
 CC moieties of glycolipids.
 CC - CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-
 CC glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-
 CC beta-D-glucosaminylglycopeptide.
 CC - CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +
 CC N-acetylglucosamine.
 CC - COFACTOR: Manganese (by similarity).
 CC - BIOPHYSICOCHEMICAL PROPERTIES:
 CC kinetic parameters:
 CC Km=63 uM for GlcNAc-B-S-pNP;
 CC - PAYWAY: Glycosylation.
 CC - SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of
 CC Golgi stack.
 CC - ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;

	CC	Isoid=060512-1; Sequence=Displayed;
	CC	Name=2;
	CC	Isoid=060512-2; Sequence=VSP_014106, VSP_014107;
	CC	Note=No experimental confirmation available;
	CC	-1- TISSUE SPECIFICITY: Found in various tissues. Highest expression
	CC	in placenta, prostate, testis, ovary, intestine and muscle, and in
	CC	fetal brain.
	CC	-1- SIMILARITY: Belongs to the glycosyltransferase 7 family.
	CC	-----
	CC	This Swiss-Prot entry is copyright. It is produced through a collaboration
	CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
	CC	the European Bioinformatics Institute. There are no restrictions on its
	CC	use as long as its content is in no way modified and this statement is not
	CC	removed.
	CC	-----
	DR	EMBL; Y12509; CAA7311.1; -, mRNA.
	DR	EMBL; AF038661; AAC39734.1; -, mRNA.
	DR	EMBL; AB024435; BAA75820.1; -, mRNA.
	DR	EMBL; AL590714; CAH72145.1; -, Genomic_DNA.
	DR	EMBL; AK023311; BAB14520.1; -, mRNA.
	DR	EMBL; BC000276; AAH00276.1; -, mRNA.
	DR	EMBL; BC006099; AAH06099.1; -, mRNA.
	DR	EMBL; BC009985; AAH09985.1; -, mRNA.
	DR	HSSP; P08037; INNM.
	DR	Ensembl; ENSG00000158650; Homo sapiens.
	DR	HGNC; HGNC:926; B4GALT3.
	DR	MIM; 604014; -.
	DR	GO; GO:0008378; F:galactosyltransferase activity; TAS.
	DR	InterPro; IPR003859; Galactosyl_T_2.
	DR	PANTHER; PTHR19300; Galactosyl_T_2; 1.
	DR	Pfam; PF02709; Galactosyl_T_2_1.
	KM	Alternative splicing; Glycoprotein; Glycosyltransferase; Golgi stack;
	KW	Manganese; Metal-binding; Multigene family; Signal-anchor;
	KM	Transferase; Transmembrane.
	FT	TOPO_DOM 1 10
	FT	TRANSMEM 11 31
	FT	Cytoplasmic (Potential).
	FT	Signal-anchor for type II membrane
	FT	protein (Potential).
	FT	Lumenal (Potential).
	FT	Manganese (By similarity).
	FT	Manganese (By similarity).
	FT	N-linked (GLCNAc...) (Potential).
	FT	N-linked (GLCNAc...) (Potential).
	FT	N-linked (GLCNAc...) (Potential).
	FT	N-linked (GLCNAc...) (Potential).
	FT	N-linked (GLCNAc...) (Potential).
	FT	By similarity.
	FT	By similarity.
	FT	PAGEBPERRAT1VPHARENNLLLYHL -> PAALPPA
	FT	PULAPACSLMHLCPEPKMKNT (in isoform 2).
	FT	/FtId=VSP_014106.
	FT	Missing (in isoform 2).
	FT	/FtId=VSP_014107.
	FT	A -> A (in Ref. 2).
	FT	G -> R (in Ref. 2).
	FT	K -> S (in Ref. 2).
	SO	SEQUENCE 393 AA; 43928 MW; CA0BF95F95ED1F4 CRC64;
	Query Match	35.3%; Score 717; DB 1; Length 393;
	Best Local Similarity	48.9%; Pred. No. 1.2e-52;
	Matches 134; Conservative	43; Mismatches 95; Indels 2; Gaps 2;
OY	95 PDNQTLOFCNQTPRLVGPRIIV-FLDEBPFKLEKIYPDTAGHGNGKDCVARHRAII	153
Dd	69 PAPQGLPCPERSPDLVGVSFSVPVSLAEIVERNRPVERPGRYRPAACEPSRTAI	128
OY	154 VPVVDREAHRIIMAHNHSLAKOOLDAITIVGVANQTENRGKLTMMVGDVASRLPYW	213
Dd	129 VPHARENNHLLLVLIHPFLORQOLAIGIVIHQAAGTNRALLLNVGVEALRDSEW	188
OY	214 QCFIFHDVLLPBDDRNLYTC-PICPIRMVAIDKFENKLPLYSIAIFGISALTYDHLKI	272
Dd	189 DCLFLHDVLDLPENDHNLYVEDCPGRPRIVANAAMKKFGSLYPPQFGVSAALTDQYLKM	248
OY	273 NGESNDFFMGWGGBDDLATRTSMAGLKVSRYPTQIARLYMTIKHSATEATPNVKCRKYIMG	332

Dd		249	NCFPNTWGWGEGDDDDATVRLAGMKSIPPPTSVCYHKNVKARGDKGNENPHRFDLLV	308
Gy		333	QTKRRWRDGLSNLKXKLVLNLEKPLVTTRAVVDL	366
Dd		309	RTQMSWTQDGNNSLTYQLARELGIVTNTITADI	342
RESULT 13				
B4GT3 PONPY				
ID B4GT3 PONPY STANDARD; PRT;		393 AA.		
AC OSNVN3:				
DT 13-SEP-2005 (Rel. 48, Created)				
DT 13-SEP-2005 (Rel. 48, Last sequence update)				
DT 13-SEP-2005 (Rel. 48, Last annotation update)				
DE Beta-1,-4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-Galtnase 3)				
DE (BetaGal-T3) [B4gal-T3] (UDP-galactose:beta-N-acetylglucosamine beta-				
-1,4-galactosyltransferase 3) (UDP-Gal-beta-GlcNAc beta-1,4-				
-galactosyltransferase 3) (Includes: N-acetyllactosamine synthase				
(EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide				
betal-1,-4-galactosyltransferase (EC 2.4.1.38); Beta-N-				
acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase				
(EC 2.4.1.1-).				
DN Name=B4GALT3;				
OS Pongo pymaeus (Orangutan).				
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;				
Pongoo.				
OX NCBI_TaxID=9600;				
[1]				
RN NOCTEOTIDE SEQUENCE [LARGE SCALE MRNA].				
RP TISSUE=Brain cortex;				
RG The German CDNA consortium;				
RU Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.				
CC -! FUNCTION: Responsible for the synthesis of complex-type N-linked				
CC oligosaccharides in many glycoproteins as well as the carbohydrate				
CC moieties of glycolipids (By similarity).				
CC -! CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-				
CC glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-				
CC beta-D-glucosaminylglycopeptide.				
CC -! CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +				
CC N-acetylglucosamine. (By similarity).				
CC -! COFACTOR: Manganese (By similarity).				
CC -! PATHWAY: Glycosylation.				
CC -! SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of				
CC Golgi stack (By similarity).				
CC -! SIMILARITY: Belongs to the glycosyltransferase 7 family.				

CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration				
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC the European Bioinformatics Institute. There are no restrictions on its				
CC use as long as its content is in no way modified and this statement is not				
CC removed.				

DR EMBL; CR925982; CAI29630.1; -, mRNA.				
DR InterPro; IPR003859; Galactosyl_T_2.				
DR PANTHER; PTRHR19300; Galactosyl_T_2; 1.				
DR Pfam; PF02709; Galactosyl_T_2; 1.				
Kw Glycoprotein; Glycosyltransferase; Golgi stack; Manganese;				
Km Metal-binding; Multigene family; Signal-anchor; Transferase;				
Tm Transmembrane.				
FT FT TOPO_DOM 1 10 Cytoplasmic (Potential). Signal-anchor for type II membrane protein (potential).				
FT FT TRANSMEM 11 31 Lumenal (Potential).				
FT FT TOPO_DOM 32 393 Manganese (By similarity).				
FT FT METAL 197 197 Manganese (By similarity).				
FT METAL 291 291 N-linked (GlcNAc...) (Potential).				
FT CARBOHYD 57 57 N-linked (GlcNAc...) (Potential).				
FT CARBOHYD 166 166 N-linked (GlcNAc...) (Potential).				
FT CARBOHYD 337 337 N-linked (GlcNAc...) (Potential).				
FT CARBOHYD 385 385 N-linked (GlcNAc...) (Potential).				
FT DISULFID 77 119 By similarity.				
FT DISULFID 190 209 By similarity.				

Seq	SEQUENCE	393 AA;	43922 MW;	CANBP9562D32D14	CRC64;
Qy	Query Match	35.3%;	Score 717;	DB 1;	Length 393;
Db	Best Local Similarity	48.9%;	Pred. No. 1,2e-52;		
Matches	134;	Conservative 43;	Mismatches 95;	Indels 2;	Gaps 2;
Qy	95 PDNQTLCRCNCTPPHNGVPIRV-FLUDEPDTFTLEKTPDTHAGHGMPKXCVARHRAII	153			
Db	69 PAPGGLPCPEPSSPLVGVSVSVSPVPSLAIEVRNDRVEPGRYRPAAGEPSPRTAII	128			
Qy	154 VPRYDREAHRLIMLNLHLSLAKOOLDPAIRIVEQVANQTFNRGKLMNVGDVASRLYPW	213			
Db	129 VPHRAREHRLFLLLYLHLHPFLQROQLAVGIVYIHQAGCFNRKALINVGRALRDREW	168			
Qy	214 QCFIFHDVLLPEDDRNLVYC-PIQPRHMSVALDKENYKLPYSALFGGISALTKDHLKI	272			
Db	189 DCLFLIHDVLLPENDHNLVYCDPCPGPRHVAANMKFGSLPYPOYFGVSALTPDYLYKM	248			
Qy	273 NGFNSDFMGCGEDDDLLATRTSMAGLKSRPTQIARVKMLKSTTEATNPNNKRYKMG	332			
Db	249 NGFNPENYMGWGEEDDIALTRVRLAGMKISRPTSGVHKVWKHGDKGNEENPHRFDLV	308			
Qy	333 QTKRRTWDGSLNKYKLVNELKLPYRAVVDL	366			
Db	309 RTQNSWTQDGMSLTYQLARELPGLYNTIADI	342			
RESULT 14	BAGT3 CRGR	STANDARD;	PRT;	395 AA.	
AC	Q80BN8;				
DT	13-SEP-2005 (Rel. 48, Created)				
DT	13-SEP-2005 (Rel. 48, Last sequence update)				
DT	13-SEP-2005 (Rel. 48, Last annotation update)				
DE	Beta-1,4-galactosyltransferase 3 (EC 2.4.1.-) (Beta-1,4-Galtrase 3)				
DE	(Beta4gal-TR3) (b4Gal-TR3) (UDP-galactose:beta-N-acetylglucosamine beta-1,4-galactosyltransferase 3) (UDP-Gal:beta-GlcNAc beta-1,4-				
DE	galactosyltransferase 3) [includes: N-acetylglucosamine synthase				
DE	(EC 2.4.1.90) (Nal synthetase); Beta-N-acetylglucosaminylglycopeptide				
DE	beta-1,4-galactosyltransferase (EC 2.4.1.38); Beta-N-				
DE	acetylglucosaminyl-glycolipid beta-1,4-galactosyltransferase				
DE	(EC 2.4.1.-)]				
GN	Name=B4GALT3;				
OS	Citricellus griseus (Chinese hamster).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;				
OC	Muridea; Cricetidae; Cricetinae; Cricetulus.				
OX	NCBI_Taxid=10029;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [MRNA].				
RC	TISSUE=Ovary;				
RA	Pubmed=14567696; DOI=10.1021/bi0353068;				
RT	Lee J., Park S.-H., Sundaram S., Raju T. S., Shaper N. L., Stanley P.;				
RT	"A mutation causing a reduced level of expression of six beta4-				
RT	galactosyltransferase genes is the basis of the Lact19 CHO				
RL	glycosylation mutant.";				
RL	Biochemistry 42:12349-12357(2003).				
CC	-1- FUNCTION: Responsible for the synthesis of complex-type N-linked				
CC	oligosaccharides in many glycoproteins as well as the carbohydrate				
CC	moieties of glycolipids (By similarity).				
CC	-1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-beta-D-				
CC	glucosaminylglycopeptide = UDP + beta-D-galactosyl-1,4-N-acetyl-				
CC	beta-D-glucosaminylglycopeptide.				
CC	-1- CATALYTIC ACTIVITY: UDP-galactose + N-acetyl-D-glucosamine = UDP +				
CC	N-acetylglucosamine. (By similarity).				
CC	-1- COFACTOR: Manganese. (By similarity).				
CC	-1- PATHWAY: Glycosylation.				
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein. Trans cisternae of				
CC	Golgi stack (By similarity).				
CC	-1- SIMILARITY: Belongs to the glycosyltransferase 7 family.				
CC	-----				
CC	This Swiss-Prot entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 18:38:45 ; Search time 6017 Seconds
(without alignments)
10883.107 Million cell updates/sec

Title: US-10-661-430-2
Perfect score: 1152
Sequence: 1 atgcgttcctgcattgcgc.....acttcacacgtgtttttag 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapexc 1.0

Searched: 5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank1:
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_scs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	1152	100.0	1152	2	AY130767	AY130767 Caenorhab
2	1152	100.0	1152	2	AY533306	AY533306 Caenorhab
3	1152	100.0	1152	2	AC025727	AC025727 Caenorhab
4	192.6	16.7	1839	6	Q0590796	Q0590796 Sequence
5	192.6	16.7	1839	6	AF132158	AF132158 Drosophila
6	192	16.7	1269	2	AY601103	AY601103 Trichopl
7	191	16.6	1189	6	AR506512	AR506512 Sequence
8	191	16.6	1879	6	AY095531	AY095531 Drosophila
9	186.8	16.2	2796	2	AB036858	AB036858 Clona int
10	181.4	15.7	2082	2	AK116794	AK116794 Clona int
11	171.8	14.9	2265	5	BC077601	BC077601 Xenopus 1
12	170.2	14.8	2681	5	BC068719	BC068719 Xenopus 1
13	169.2	14.7	1338	9	AY117537	AY117537 Cricetulu
14	168.4	14.6	2007	4	BC061812	BC061812 Rattus no
15	168.2	14.6	1921	4	BT020916	BT020916 Bos tauru
16	168.2	14.6	1927	4	BT020682	BT020682 Bos tauru
17	168.2	14.6	1949	4	BT020868	BT020868 Bos tauru
18	167.6	14.5	1182	8	HSUDP GAL	Y12509 Homo sapien

19	167.6	14.5	1193	8	AB024435	AB024435 Homo sapi
20	167.6	14.5	1912	8	BC009985	BC009985 Homo sapi
21	167.6	14.5	1934	6	BC006099	BC006099 Homo sapi
22	167.6	14.5	1938	6	BD206381	BD206381 Human nuc
23	167.6	14.5	1938	6	AR400696	AR400696 Sequence
24	167.6	14.5	1938	6	AX013200	AX013200 Sequence
25	166.4	14.4	4819	5	AJ851705	AJ851705 Gallus ga
26	166	14.4	1885	8	AK092493	AK092493 Homo sapi
27	166	14.4	1915	8	BC000276	BC000276 Homo sapi
28	166	14.4	1920	8	AP038661	AP038661 Homo sapi
29	166	14.4	2340	5	BC075452	BC075452 Xenopus t
30	165.8	14.4	1954	5	CR760293	CR760293 Xenopus t
31	164.4	14.3	2479	2	AK112685	AK112685 Clona int
32	163.2	14.2	2626	5	BC041742	BC041742 Xenopus 1
33	161.2	14.0	2147	8	AF069054	AF069054 Homo sapi
34	161.2	14.0	2289	8	AP038664	AP038664 Homo sapi
35	161.2	14.0	2663	8	CR26530	CR26530 Xenopus t
36	159.6	13.9	1177	8	AB024742	AB024742 Homo sapi
37	159.6	13.9	1198	8	BC074835	BC074835 Homo sapi
38	159.6	13.9	1198	8	BC074884	BC074884 Homo sapi
39	159.6	13.9	1329	5	BC087436	BC087436 Xenopus 1
40	159.6	13.9	1355	8	BC069642	BC069642 Homo sapi
41	159.6	13.9	3931	8	AF097159	AF097159 Homo sapi
42	158	13.7	1962	9	BC013619	BC013619 Mus muscu
43	157.4	13.7	2098	5	CR942597	CR942597 Xenopus t
44	157.2	13.6	1853	6	BD191505	BD191505 Secrete
45	157	13.6	1714	8	BC004523	BC004523 Homo sapi

ALIGNMENTS

RESULT 1
AY130767 1152 bp mRNA linear INV 17-SEP-2002
LOCUS
DEFINITION Caenorhabditis elegans UDPgalNAc:GlcNAc[beta]-R
{beta}1,4-N-acetylglucosaminyltransferase mRNA, complete cds.
ACCESSION
AY130767
VERSION
AY130767.1 GI:22415754

KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans
Caenorhabditis elegans
Bukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE
1 (bases 1 to 1152)
Kawar, Z.S., Van Die, I. and Cummings, R.D.
Molecular Cloning and Enzymatic Characterization of a
UDP-GalNAc:GlcNAc[beta]-R beta 1,4-N-Acetylglucosaminyltransferase
from Caenorhabditis elegans

JOURNAL
J. Biol. Chem. 277 (38), 34924-34932 (2002)

PUBMED
2 (bases 1 to 1152)
Kawar, Z.S., Van Die, I. and Cummings, R.D.
Direct Submission
Submitted (10-JUN-2002) Biochemistry and Molecular Biology,
University of Oklahoma Health Sciences Center, 975 NE 10th St. BRC
Rm. 417, Oklahoma City, OK 73104, USA

FEATURES
source
1..1152
Location/Qualifiers

CDS
1..1152
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/note="Ce[beta]4GalNAcT, similar to Caenorhabditis elegans
hypochemical protein Y7357A.7"
/codon_start=1
/product="UDPgalNAc:GlcNAc[beta]-R
{beta}1,4-N-acetylglucosaminyltransferase"
/protein_id="AA095168.1"
/db_xref="GI:22415754"
/translation="MAFRHLAVARIKSLVLCVALLVHAMITKPSLYENITTSST
LIADVAMEAVLGNASTSDLLDTWNTFSISVNTGSEMEDRPLIFPDNQLQF
CNQTPPHLVGPIRVLPDPPFKTEKIVPDTAGHGMPKDCVARRVAIVPYDRE

AHLRLMLHNLHSLAKQOLDYALFIVEOVANOTFNKLMNVGDVASHLYPWOCFIF
HDVDLIPEDDRNLYTCPIOPRHMSVAIDENKYLPSYALFGISALTXHKLKINFIS
NDPMWGGEDDILATRTSMAGLKVSRYPQIARVKMIKSTETATNVNCRXYKIMGQT
KRRWTDGJSLNKKYKLVNLEKPLVTRAVVDLLEKDCRRLRDPFTCC"

ORIGIN

Query Match 100.0%; Score 1152; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTTTGTCAATTTGGCAGTCGCCAGACTCAAGTCCTGCTCGTACTTTGTGCCGCTT 60
DB 1 ATGGCTTTTGTCAATTTGGCAGTCGCCAGACTCAAGTCCTGCTCGTACTTTGTGCCGCTT 60
QY 61 CTTCATTAAGTTCATGCAATGATTTATTAAGTTCATGCTGCTTTACGAGAACTTACTATC 120
DB 61 CTTCATTAAGTTCATGCAATGATTTATTAAGTTCATGCTGCTTTACGAGAACTTACTATC 120
QY 121 GGCTCCCTCGACCTTATTTGCGACGTCGACGCAATGAGGCAAGTGTGCGGAAATACGGCT 180
DB 121 GGCTCCCTCGACCTTATTTGCGACGTCGACGCAATGAGGCAAGTGTGCGGAAATACGGCT 180
QY 181 TCACATTCGGATGATCTACTTGTGATCGTGAATTCACGTTTTCACCGATTTCTGAAGTT 240
DB 181 TCACATTCGGATGATCTACTTGTGATCGTGAATTCACGTTTTCACCGATTTCTGAAGTT 240
QY 241 AATCAGACTAGTTTATGAGGAGCATTCGTCATCCGTTCCCGGCAACAGACTCTT 300
DB 241 AATCAGACTAGTTTATGAGGAGCATTCGTCATCCGTTCCCGGCAACAGACTCTT 300
QY 301 CAATTCGTATGACAGACACCTCCCACTGTCGAGCCCAATCCGTGATTCCTCGATGAG 360
DB 301 CAATTCGTATGACAGACACCTCCCACTGTCGAGCCCAATCCGTGATTCCTCGATGAG 360
QY 361 CCGGACTTCAAACTCTCGAGAAATTTATCCGACACGACGCGGTGACATGGAATG 420
DB 361 CCGGACTTCAAACTCTCGAGAAATTTATCCGACACGACGCGGTGACATGGAATG 420
QY 421 CCTAAGGATGTGTGCAAGGCAATCGTGTGCTATTTATGTCGCTTATAGATCGGAA 480
DB 421 CCTAAGGATGTGTGCAAGGCAATCGTGTGCTATTTATGTCGCTTATAGATCGGAA 480
QY 481 GCAATTTGAGAAATATGCTCCACAAATTTGCACTGTTGTCGCAAAACAATTTGGAC 540
DB 481 GCAATTTGAGAAATATGCTCCACAAATTTGCACTGTTGTCGCAAAACAATTTGGAC 540
QY 541 TATGCAATTTTCAATTTGAGCAAGTGGCAATCGATCGATTCGCGGAAACTATATG 600
DB 541 TATGCAATTTTCAATTTGAGCAAGTGGCAATCGATCGATTCGCGGAAACTATATG 600
QY 601 AACGTTGATACGACGTAAGATCAAGCTCTAACCCATGGGAGGCTTCATCTTTCATGAT 660
DB 601 AACGTTGATACGACGTAAGATCAAGCTCTAACCCATGGGAGGCTTCATCTTTCATGAT 660
QY 661 GTGATTTAAGTCCCGGAAGATGACCGTAACTGTACAGTGTCCAAATTCACCAACGTCAT 720
DB 661 GTGATTTAAGTCCCGGAAGATGACCGTAACTGTACAGTGTCCAAATTCACCAACGTCAT 720
QY 721 ATGAGTGTAGCGATCGATTAATTCATTTAACTTTCATATTCGCGGAGCTTCGCGGGA 780
DB 721 ATGAGTGTAGCGATCGATTAATTCATTTAACTTTCATATTCGCGGAGCTTCGCGGGA 780
QY 781 ATGAGTGTAGCGATCGATTAATTCATTTAACTTTCATATTCGCGGAGCTTCGCGGGA 840
DB 781 ATGAGTGTAGCGATCGATTAATTCATTTAACTTTCATATTCGCGGAGCTTCGCGGGA 840
QY 841 GGTGGGGCGAGAGAGAGAGATTTGGCGAGCAATCGATGAGTGGCTGCACTGAAAGTT 900
DB 841 GGTGGGGCGAGAGAGAGAGATTTGGCGAGCAATCGATGAGTGGCTGCACTGAAAGTT 900
QY 901 TCAAGATATCCGACAAATTTGCAAGATATTAAGCTGATTAAGCACTGACGGAAGCGACG 960
DB 901 TCAAGATATCCGACAAATTTGCAAGATATTAAGCTGATTAAGCACTGACGGAAGCGACG 960

QY 961 AATCCAGTTAATAATGCCGCTACAAAAATATATGAGCCAAACGAGGCCGATGACACGT 1020
DB 961 AATCCAGTTAATAATGCCGCTACAAAAATATATGAGCCAAACGAGGCCGATGACACGT 1020
QY 1021 GACGCGCTAAGCAATCTGAATTAAGCTGCTGAATCTGGAATTTGAAGCTCTTACACT 1080
DB 1021 GACGCGCTAAGCAATCTGAATTAAGCTGCTGAATCTGGAATTTGAAGCTCTTACACT 1080
QY 1081 CGAGCGCGTGTGATTTGCTCGAAAAAGACTCGCCCGGAGACTGTGGAAGGACTTTCCA 1140
DB 1081 CGAGCGCGTGTGATTTGCTCGAAAAAGACTCGCCCGGAGACTGTGGAAGGACTTTCCA 1140
QY 1141 ACGTGTTTTAG 1152
DB 1141 ACGTGTTTTAG 1152

RESULT 2
AY533306 1152 bp mRNA linear INV 23-FEB-2004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Pelodetinae; Caenorhabditis.

REFERENCE
AUTHORS
1 (bases 1 to 1152)
Griffiths,J.S., Huffman,D.L., Whitacre,J.L., Barrows,B.D.,
Maroquin,L.D., Muller,R., Brown,J.R., Henmet,T., Esko,J.D. and
Aroian,R.V.
Resistance to a bacterial toxin is mediated by removal of a
conserved glycosylation pathway required for toxin-host
interactions
J. Biol. Chem. 278 (46), 45594-45602 (2003)

JOURNAL
PUBMED
12944392
REFERENCES
AUTHORS
TITLE
JOURNAL
Submitted (26-JAN-2004) Biology, Univ. California, San Diego, 9500
Gilman Dr, La Jolla, CA 92093, USA
Location/Qualifiers

FEATURES

source
1..1152
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/chromosome="I"
/map="Y73E7A"
1..1152
/gene="bre-4"
1..1152
/gene="bre-4"
/gene="bre-4"
/note="glycosyltransferase; required for susceptibility to
Bacillus thuringiensis Cry toxins"
/codon_start=1
/product="BRE-4"
/protein_id="AA521308.1"
/db_xref="GI:42601222"
/translation="MAFRLAVARLKLILVCAVLLVHAMTKLPSLYENLTIGSST
LIADVAMEAVLGNFTASTDDLDLTWNSFTSPISVNOQTSFMDIARPIIFPDNQLQF
CNOTPHLVGPTRVFLDEDPFTLEKIYDTHAGHGMPKDCVABRVAIIVYDRDE
AHLRLMLHNLHSLAKQOLDYALFIVEOVANOTFNKLMNVGDVASHLYPWOCFIF
HDVDLIPEDDRNLYTCPIOPRHMSVAIDENKYLPSYALFGISALTXHKLKINFIS
NDPMWGGEDDILATRTSMAGLKVSRYPQIARVKMIKSTETATNVNCRXYKIMGQT
KRRWTDGJSLNKKYKLVNLEKPLVTRAVVDLLEKDCRRLRDPFTCC"

ORIGIN

Query Match 100.0%; Score 1152; DB 2; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGGCTTTTCGATCTATTGGCAGTCGCCAGACTCAAGTCGTGCTGCTACTCTTGTCGGCT 60
DB 1 ATGGCTTTTCGATCTATTGGCAGTCGCCAGACTCAAGTCGTGCTGCTACTCTTGTCGGCT 60
QY 61 CTTTCTATTGTTGATGCAATGATTTATTAAGATTCCATGCTTTAGAGAACTTACTATC 120
DB 61 CTTTCTATTGTTGATGCAATGATTTATTAAGATTCCATGCTTTAGAGAACTTACTATC 120
QY 121 GGCTCCTCGACCCCTTATTCGCCGCGTGCAGCGCAATGGAGGAGTGTCTGGGAAATACGGCT 180
DB 121 GGCTCCTCGACCCCTTATTCGCCGCGTGCAGCGCAATGGAGGAGTGTCTGGGAAATACGGCT 180
QY 181 TCCACTTCCGAGTATCTTATGATAGTGAATTCACGCTTTTCCAGATTTCTGAGTT 240
DB 181 TCCACTTCCGAGTATCTTATGATAGTGAATTCACGCTTTTCCAGATTTCTGAGTT 240
QY 241 AATCAGACTAGTTTATGAGAGACATTCCTCCATCTGTTCCCGCAACACGACTCTT 300
DB 241 AATCAGACTAGTTTATGAGAGACATTCCTCCATCTGTTCCCGCAACACGACTCTT 300
QY 301 CAATTCTGTAATCAGACCTCCCGACCTCGTGCAGCCATCCGCTATTTCTCGATGAG 360
DB 301 CAATTCTGTAATCAGACCTCCCGACCTCGTGCAGCCATCCGCTATTTCTCGATGAG 360
QY 361 CCCGACTTCAAACTCTCGAGAAATCTATCCGAGACGCGCAGCGTGAATGAGATG 420
DB 361 CCCGACTTCAAACTCTCGAGAAATCTATCCGAGACGCGCAGCGTGAATGAGATG 420
QY 421 CCTAAGGATTTGTTGCAAGGCAATCGTGTCTATTTATTTGTCCTTATAGATCTGAA 480
DB 421 CCTAAGGATTTGTTGCAAGGCAATCGTGTCTATTTATTTGTCCTTATAGATCTGAA 480
QY 481 GGAACATTTGAGAAATATGTCACAAATTTGCACTGCTGCTGCCAAACAAATTTGAGC 540
DB 481 GGAACATTTGAGAAATATGTCACAAATTTGCACTGCTGCTGCCAAACAAATTTGAGC 540
QY 541 TATGCAATTTTCAATTTGAGAGCAAGTGGCAATCAGACCTTTAATGCGGGAAACTAATG 600
DB 541 TATGCAATTTTCAATTTGAGAGCAAGTGGCAATCAGACCTTTAATGCGGGAAACTAATG 600
QY 601 AACGTTGATCAGACGATGATCAGCTCTTACCCATGSCAGTCTTCAATCTTTCAATGAT 660
DB 601 AACGTTGATCAGACGATGATCAGCTCTTACCCATGSCAGTCTTCAATCTTTCAATGAT 660
QY 661 GTGATTTTCTGCCCCGAAGATGACCTTAACCTGTAACGTCGTCATTTCAACCACTCAT 720
DB 661 GTGATTTTCTGCCCCGAAGATGACCTTAACCTGTAACGTCGTCATTTCAACCACTCAT 720
QY 721 ATGAGTGTAGCATGATTAATTTCAATTTAACTTCCATTTTGGCGATCTTGGCGGA 780
DB 721 ATGAGTGTAGCATGATTAATTTCAATTTAACTTCCATTTTGGCGATCTTGGCGGA 780
QY 781 ATCAGTGCATACAAAGATCAGCTGAAGAAATCAATGATTTTGAATGATTTTGG 840
DB 781 ATCAGTGCATACAAAGATCAGCTGAAGAAATCAATGATTTTGAATGATTTTGG 840
QY 841 GGTGGGGGGGAGAGACGATTTGGGAGAGAACTGATGCTGTGACCTGAAAGTT 900
DB 841 GGTGGGGGGGAGAGACGATTTGGGAGAGAACTGATGCTGTGACCTGAAAGTT 900
QY 901 TCAAGATATCCGACAAATTTGACATATTAATTAATTAATTAATTAATTAATTAATTAAT 960
DB 901 TCAAGATATCCGACAAATTTGACATATTAATTAATTAATTAATTAATTAATTAATTAAT 960
QY 961 AATCCAGTTAATAATGCGCTACAAATTAATGGGCAACGAAAGGCGGATGAGACAGT 1020
DB 961 AATCCAGTTAATAATGCGCTACAAATTAATGGGCAACGAAAGGCGGATGAGACAGT 1020
QY 1021 GACGGCTTAAGCAATGTAAGTAACTGTAATCTGTAATCTGTAATCTGTAATCTGTA 1080
DB 1021 GACGGCTTAAGCAATGTAAGTAACTGTAATCTGTAATCTGTAATCTGTAATCTGTA 1080
QY 1081 CCAGCGCTGTGCTGCGAATAAGACTGCGCGGAGAGCTGGAAGGAGCTTTCCA 1140
```

```
DB 1081 CCAGCGCTGTGCTGCGAATAAGACTGCGCGGAGAGCTGGAAGGAGCTTTCCA 1140
QY 1141 ACGTGTTTTAG 1152
DB 1141 ACGTGTTTTAG 1152

RESULT 3
AC025727/c 63747 bp DNA linear INV 27-MAY-2005
LOCUS AC025727
DEFINITION Caenorhabditis elegans cosmid Y73E7A, complete sequence.
ACCESSION AC025727
VERSION AC025727.3 GI:14530915
KEYWORDS HTG.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita;
Rhabdilitida; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 63747)
2 (bases 1 to 63747)
3 The sequence of C. elegans cosmid Y73E7A
4 (bases 1 to 63747)
5 (bases 1 to 63747)
6 (bases 1 to 63747)
7 (bases 1 to 63747)
8 (bases 1 to 63747)
9 (bases 1 to 63747)
10 (bases 1 to 63747)

REFERENCE
AUTHORS Du,H. and Maupin,R.
TITLE The sequence of C. elegans cosmid Y73E7A
JOURNAL Unpublished (2001)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 63747)
5 (bases 1 to 63747)
6 (bases 1 to 63747)
7 (bases 1 to 63747)
8 (bases 1 to 63747)
9 (bases 1 to 63747)
10 (bases 1 to 63747)

REFERENCE
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
7 (bases 1 to 63747)
8 (bases 1 to 63747)
9 (bases 1 to 63747)
10 (bases 1 to 63747)

REFERENCE
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
9 (bases 1 to 63747)
10 (bases 1 to 63747)

REFERENCE
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
10 (bases 1 to 63747)

REFERENCE
AUTHORS
```

CONSRM WormBase Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE
AUTHORS 11 (bases 1 to 63747)
TITLE WormBase Consortium
JOURNAL Direct Submission
Submitted (05-MAR-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

REFERENCE
AUTHORS 12 (bases 1 to 63747)
TITLE WormBase Consortium
JOURNAL Direct Submission
Submitted (27-MAY-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

COMMENT
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: submissions@watson.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m3 subclone.

For a graphical representation of this clone sequence and its analysis see:
<http://www.wormbase.org/db/seq/sequence?name=Y73E7A;class=sequence>

NEIGHBORING CLONE INFORMATION

The 5' clone is Y92H12A, 200 bp overlap; the 3' clone is Y71G12B, 4000 bp overlap.

NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program GeneIndex (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFome cloning project (<http://worfdb.dcfi.harvard.edu/>), similarity to other proteins from Blastx analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. rRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES
source
1.63747
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="Y73E7A"
5487..12773

CDS
/locus_tag="Y73E7A.3"
join(5487..5659,7957..8063,8778..8971,9035..9119,10020..10254,11408..11615,11668..11742,11836..11943,12669..12773)
/locus_tag="Y73E7A.3"
/standard_name="Y73E7A.3"
/note="contains similarity to Pfam domain PF06027 (Bukaryotic protein of unknown function (DUF914)); coded for by the following C. elegans cDNAs: YK1142A05.3, OSTF3045C02"
/codon_start=1
/product="Hypothetical protein Y73E7A.3"
/protein_id="AAG23391.3"
/db_xref="GI:60543518"
/db_xref="WormBase:Y73E7A.3"
/translation="MTNCGAESTSAPQANGHLRPPGIOPHRNPTPLPSPDEPD AINCPCCDDPVRVSTFRKALLIGQLISLCTCGTSSQLNQNNAAPQAFSRYFL LCFVTCISLACNDENGLVVRKRGKRWLIIAIDVEANITVAIYOTNLTSVQL DCAITPVYFLFSLFSLVRYLASHITIGVITCIIGACVYADALDQKALDQGSNKL GDIICLAAMVAICNVAAEELVKOHSREYVIGLFGCIYGVQTAFAVEQALSKI VMDGTIVSYFALFARSMFLFSLVTVLOKTSALMFINSLTADFSLLFGIFMFKDT FHYLFEVSRIVCTIGSVTYSKERTQVRADDERRVCPCLFVCCCGCCCFEBGDSREG SIDVSPSPNERQKGLNPRGSMSPCPVHGSQQLSHL"
complement(14300..15495)
/locus_tag="Y73E7A.4"
complement(join(14300..14457,15222..15495))
/locus_tag="Y73E7A.4"
/standard_name="Y73E7A.4"
/note="contains similarity to Pfam domain PF05835 (Synaphin protein); coded for by the following C. elegans cDNAs: YK1215A10.5, YK1728H08.3, OSTF30035H07"
/codon_start=1
/product="Hypothetical protein Y73E7A.4"
/protein_id="AAG23387.1"
/db_xref="GI:10800392"
/db_xref="WormBase:Y73E7A.4"
/translation="MAGFLMKQVNOQSEVTGIGLGMKDDGSEKTEGDEPVIAARL EOEERKEKRGKRNENRERKROGIDKVIKKKEGVMADFPBGRIGPRKTPETIAA ENNAEDDSLIGQLGTEQVAKATWATGAFETVKGFPFPGK"
complement(19944..20816)
/locus_tag="Y73E7A.5"
complement(join(19944..20130,20270..20541,20589..20816))
/locus_tag="Y73E7A.5"
/standard_name="Y73E7A.5"
/note="contains similarity to Pfam domain PF07735 (P-box associated)"
/codon_start=1
/product="Hypothetical protein Y73E7A.5"
/protein_id="AAG23386.1"
/db_xref="GI:10800391"
/db_xref="WormBase:Y73E7A.5"
/translation="MFOHIFIKFOIIGDIFRLSYOSALSHPLFRKCHFLKFLGEGSYI TTBDMKFTIKNNKIKGPTFRSLRKLIDIKLVNIPRLDPEKSPGLTLDLKMNCE KRFEDLNTIKAREINKYIRVWDGDLPLRLQLNYYHMAAGLLVLPYSKNNPK NBARLPCDLEFVDCSYGRDIERSDGLATILLENFAMFVLVHDFPDPRTLPHFS NNSVSPITNN"
21160..26092
/locus_tag="Y73E7A.2"
join(21160..21238,21285..21638,23801..24363,25697..26092)
/locus_tag="Y73E7A.2"
/standard_name="Y73E7A.2"
/note="contains similarity to Rattus norvegicus AATP protein (Apoptosis antagonizing transcription factor); TR:Q9QYW0; coded for by the following C. elegans cDNAs:

Query Match 22.7%; Score 262; DB 2; Length 63747;
Best local Similarity 100.0%; Pred. No. 2e-68; Indels 0; Gaps 0;
Matches 262; Conservative 0; Mismatches 0;

QY 1 ATGGCTTTGCGATTGGCAGTGGCCAGACGTAAGTCGTTGCTACTTTGCGCGTT 60
DB 45925 ATGGCTTTGCGATTGGCAGTGGCCAGACGTAAGTCGTTGCTACTTTGCGCGTT 45866

Oy		61	CTTCAATAGTTGCAATGCATATATTATAAAGATTCATCGCTTTACGAACAACCTTAACATC	120
Dd		45865	CTTCATTAATGTTCAATGCAATATATTATAAGATTCATCGCTTTACGAACAACCTTAACATC	45808
Oy		121	GAGTCTTGACCCCTTATTTGCCGACGTGACGCCAATGAGGAGCATGCTCGGAAATACGGCT	180
Dd		45805	GGCTCTCGAACCCCTTATTTGCCGACGTGACGCCAATGAGGAGCATGCTCGGAAATACGGCT	45744
Oy		181	TCCAATTGGATNGATCTACTTGATACGTGGGAATTCACAGTTTTTCAACGATTTTCGAAGTT	240
Dd		45745	TCCAATTGGATNGATCTACTTGATACGTGGGAATTCACAGTTTTTCAACGATTTTCGAAGTT	45688
Oy		241	AATCAGACTAGTTTATGAGAG 262	
Dd		45685	AATCAGACTAGTTTATGAGAG 45664	
<hr/>				
RESULT 4				
COS90796				
LOCUS	COS90796	1839 bp	DNA	linear PAT 02-FEB-2004
DEFINITION	Sequence 18554 from Patent WO0171042.			
ACCESSION	COS90796			
VERSION	COS90796.1	GI:41649177		
KEYWORDS	.			
SOURCE	Drosophila sp.			
ORGANISM	Drosophila sp. Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peirygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Empidoidea; Drosophilidae; Drosophila.			
REFERENCE	1 Venter,J.C., Adams,M., Li,P.W. and Myers,E.W. Detection Kites, such as nucleic acid arrays, for detecting the expression of 10,000 or more Drosophila genes and uses thereof Patent: WO 0171042-A 18554 27-SEP-2001; PE Corporation (NY) (US)			
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source	location/Qualifiers 1..1839 /organism="Drosophila sp." /mol_type="unassigned DNA" /db_xref="taxon:7242"			
ORIGIN				
Query Match	16.7%; Score 192.6; DB 6; Length 1839;			
Best Local Similarity	58.7%; Pred. No. 3e-47;			
Matches 333; Conservative	0; Mismatches 234; Indels 0; Gaps 0;			
Oy		399	GCAGCGCGRGGACATGATAGCCCAAAGATGTGTGGCAAGCATCGTGTGCTATATAT	458
Dd		600	GCGCCCTGTGTGCGCCCTTGAGACCCTGAAGAATCGCAATCCAGCATCACGTGGCTATTGT	659
Oy		459	TGTGCCCTATAGAGATCGTGAAGCATTTGAGAATAATGCTCCACAAATTTGCACTCGTT	518
Dd		660	TGTGCCCTCGCGCATGATAGGCCCATCTMTACTTTTCTGTGGCAACATCCACCATT	719
Oy		519	GCTGCCCAACAAACAATTGACATATGCAATTTTCAATTTGTGAGCAAGTGGCGAATCAGAC	578
Dd		720	TCTGTAGAAGACGCGCATCGCCCTATCGCATTTTCAATGTAGAGACAAACGGAAGCC	779
Oy		579	GTTTAATTCGGCGGAAACTTAAGAAAGTTGGATAGCAGCTAGCATCAGCCCTCTACCATG	638
Dd		780	CTTTAATTCGGCGCTGCGCATGATGAACATTTGGTATTTTGGAGGCGCTTAAAGCTGTACAGTG	839
Oy		639	GCAGTCTTCATCTTTCATGATGTGCAATTTTACGCGCGGAAGATGACCGTAACTGTACAC	698
Dd		840	GGATTTGTTTATATTCACAGATGTGCAATCTTCTGTGCTTTGAGCAGCGCAATCTTTACAA	899
Oy		699	GTTTCCAAATTCACACGTCATATGAGTGTAGCGATGATTAATTCATTAATAACTTCC	758
Dd		900	CTGTTCACGTGACGCGCACACATGTCAGTGGCTATATAGACACGCTGAACCTTCAAGTTGCC	959
Oy		759	ATATTTGGCGCATTTTTCGGCGGAATACGTGCACTAAACAAAAGATCACTCGGAAGAAATCA	818
Dd		960	TTATTCATCAATATTTGGAGGTGTTTTCGCAATGACGGTGTAGCACTTTTCAGGCGGTAA	1019

OY	819	TGATTTTTCGAATATTTTTGGAGGTGGCGGAGAGACACAAGATTGGGAGCAGAAC	878
DB	1020	TGGATTCTCAAACTCCGTTCTTTGGCTGGGGCGGAGAGATGACAKATGTCCACAGGTT	1079
OY	879	ATCGATGCCTGCAGTGAAGAGTTTCAGATATCCGACACAAATTGCACCATATAAATGAT	938
DB	1080	GAAACACGCCAACCTATTATCATATCAAGATATCCGGTGCAACTAGCTCGTCAAGATGCT	1139
OY	939	TAAGCACTCCAGCGAAGCGAAGATCC	965
DB	1140	GAAACATTCAGAGAAAAGGCCAATCC	1166
RESULT 5			
LOCUS	AF132158	2269 bp	mRNA linear INV 03-JUN-1999
DEFINITION	Drosophila melanogaster clone GH13356 unknown mRNA.		
VERSION	AF132158		
KEYWORDS	AF132158.1 GI:4972701		
SOURCE	FLI CDNA.		
ORGANISM	Drosophila melanogaster (fruit fly)		
REFERENCE	Drosophila melanogaster Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 2269) Rubin,G.M., Wan,K.H., Harvey,D., Lewis,S.E., Brokstein,P., Tsang,G., Aggayanl,A., Arcaina,T.T., Baxter,B., Blazer,R.G., Butenoff,C., Champe,M., Chavez,C., Chew,M., Doyle,C.M., Fafian,D.E., Friese,E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Evans-Holm,M., Houston,K.A., Hummasti,S.R., Kim,E., Li,P., Moshrefi,M., Pacleb,J.M., Park,S., Segreita,A., Seth,H., Snir,E., Svirkas,R.R., Weinburg,T. and Celniker,S.E. Direct Submission Submitted (01-MAR-1999) Berkeley Drosophila Genome Project, University of California Berkeley, 539 Life Sciences Addition 3200, Berkeley, CA 94720, USA Sequence submitted by: Berkeley Drosophila Genome Project University of California Berkeley Berkeley, CA 94720 For further information about this sequence,including its location and relationship to other sequences, please visit our sequence archive Web site (http://fruitfly.berkeley.edu/sequence/) or send email to cdaa@fruitfly.berkeley.edu . Location/Qualifiers 1..2269 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /cldone="GH13356" 519..1730 /codon_start=1 /product="unknown" /protein_id="AAD34746.1" /db_xref="GI:4972702" /db_xref="FlyBase:FBCD0013356" /translatation="MIDFTLANLIRFLAGICLLLVNVEFGNSDGGSATSLSKLSLR RVHYKTHAYGNASDGAAGSEASRLPSALSDRDRDLNGGPSTRTVATVAN FTSPDILTRFLGLTKFKLPKROKSTALANCDDPRDGGPTTPMTTESLIVBA ELGELRGGAEPENCNOAHVAIVVFEDRYAHLILPLRNHPFLMKORIAVRIFI VEONGENPRAAMANGTYEALXIKWCMDFITHDVLFLDNBNLCROPRHMSVV AIDTNTRLPYRSIFPGVSAMTREHQAUNVFNSSPKMGEGEDDMENRLKHNLPTS RYPNNTARIOMLKHKQEKANPKXYENLDNGMSKI BDGDJINSIKTSITSITQOFPFTTWY LAELKNERSK"		
FEATURES			
SOURCE			
CDS			
COMMENT			
JOURNAL			
TITLE			
ORIGIN			
Query Match	16.7%	Score 192.6;	DB 2; Length 2269;
Best Local Similarity	58.7%	Pred. No. 3,1e-47;	
Matches 333;	Conservative 0;	Mismatches 234;	Indels 0; Gaps 0;
399	GCAAGCGGTGAGCATGAATGCTTAAGATGTGTGTGCAAGGATCGTGTCTATAT	458	

Db 1016 GGGCCCTGGTGGGCGCTTTCAGAGCTGAAAACTGCATATGCCAGCATCAGTGGCTATTGT 1075
Qy 459 TGTGCCCCTATAGAGATCGTGAAGCACAATTGGAATAAGTCTCCACAAATTGGACCTGTT 518
Db 1076 TGTGCCCCTTCGGGATCGATACGCCCATCTATTACTTCTTCGGCAACAATCCACCATTT 1135
Qy 519 GCTCGCCAAACAACAATTGGAATGATGCAATTTTCATTGTGAGCAAGTGGCGCATCGAC 578
Db 1136 TCTGATGAAGCAGCGGATGGCCATATGCAATTTTCATTGTAGACACACCAAGCGGAAAGCC 1195
Qy 579 GTTTAAATCGCGGAAACTAATGACCTTGATACGACGTACATCAGCCTTCAATG 638
Db 1196 CTTTAAATCGGCGCTGCATGATGAACAATGGTATTGTGAGGCGCTTAAAGCTGTACAGTG 1255
Qy 639 GCAAGCTTCATCTTTCATGATGATGATTTTACCTGCCCAAGATGACCGTAACTGTACAC 698
Db 1256 GGAATGTTTAAATTCACAGATGTGATCTTCTGCTTGGACGACCGCAATCTTTTACAA 1315
Qy 699 GTGTCCAAATTCACAACAGTCATATGATGATGATGATGATGATGATGATGATGATGATG 758
Db 1316 CTGTCCACGTGACGCGCGGACACATGTCAAGTGTATAGACAGCTGATTCAGTGTGCG 1375
Qy 759 AATATTCGGGATCTTTCGGCGGAAATCAGTGCATTAACAAAATCAGCTGAAATCAA 818
Db 1376 TTAATCATCAATATTGTGAGGTGTTTCCGCAATGACGCGTGAACATTTCAAGCGCTGAAA 1435
Qy 819 TGGATTTTGAATGATTTTGGGTTGGGCGGAGAGAGACAGATTTGGCGAGAAC 878
Db 1436 TGGGTTCTCAAACTGTTCTTTGGCGGCGGCGAGATGACAGATGTCACACAGTTT 1495
Qy 879 ATCGATGCTGCTGACTGAAGTTTCAAGATATCCGACACAATTTGCAAGATATTAATGAT 938
Db 1496 GAAGCAGCGAAACTATTTCATATCAAGGTATCCGGTCAACATAGCTCGTACAAAGATGCT 1555
Qy 939 TAAAGCATCGACGGAAGCGACGAATCC 965
Db 1556 GAAGCATCAGAAAGAAAGCGCAATCC 1582

RESULT 6
AY601103 1269 bp mRNA linear INV 02-AUG-2004
LOCUS Trichoplusia ni beta 1,4-N-acetylglucosaminyltransferase mRNA,
DEFINITION complete cds.
ACCESSION AY601103
VERSION AY601103.1 GI:47156062
KEYWORDS Trichoplusia ni (cabbage looper)
SOURCE Trichoplusia ni
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Plusiinae; Trichoplusia.
REFERENCE 1 (bases 1 to 1269)
AUTHORS Vadate,N. and Jarvis,D.L.
TITLE Molecular Cloning and Functional Characterization of a Lepidopteran
Insect [beta]4-N-Acetylglucosaminyltransferase with Broad
Substrate Specificity, a Functional Role in Glycoprotein
Biosynthesis, and a Potential Functional Role in Glycolipid
Biosynthesis
J. Biol. Chem. 279 (32), 33501-33518 (2004)
PUBMED 15173167
REFERENCE 2 (bases 1 to 1269)
AUTHORS Vadate,N. and Jarvis,D.L.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2004) Molecular Biology, University of Wyoming,
Laramie, WY 82071, USA
FEATURES
source
1. 1269
/organism="Trichoplusia ni"
/mol_type="mRNA"
/db_xref="taxon:7111"
1. 1266

/note="b4galNACT; glycoeyltransferase"
/codon_start=1
/product="beta 1,4-N-acetylglucosaminyltransferase"
/protein_id="AA11926.1"
/db_xref="GI:47156063"
/translation="WGGRATRAARLLLLVLAALAAVEYLFSGIILASPLRTYLPYLY
NATOPTLRNVERLANWPKIIPSNITLSEESIISINLSNHTTRASVHPSSITET
ASKLDKNNTIDGAPAMISPTPLITTKMDSIKSVTTEDGVKAELAVVTLPLCDMP
PDIGITLNTKTELELEWTEKKEPEVEWGGYRSPNCTARHRAIIVPRDROHLAIF
LNHMRFLMKQOIEYGI FIVEOGKNDENRAKLVNMGVESQKLVAEQMOCVFPHID
LLPLDRUYGCPQPRHMSASIDLHRLKPYEDI FGVSAVMTLEQFTRVNGFSNRYW
GHEGDDMSYRIKINHYIARIYKRSIARIYALMDHKSTPNPKRYQLLSQTSKTRQXD
GLSTLEYELVQVQYHLYTHILVINDERS"

ORIGIN
Query Match 16.7%; Score 192; DB 2; Length 1269;
Best Local Similarity 58.3%; Pred. No. 4.4e-47;
Matches 358; Conservative 0; Mismatches 250; Indels 6; Gaps 1;
Qy 358 GAGCCCGACTTGCAAACTCTCGAGAAATCTATCCGACACGACCGCGGTGACATGGA 417
Db 514 GAACTGGAACCTGGAATGGGTGAGAAAAAGTTCCGAGAGTGAAGTGGGAGCGGCTAC 573
Qy 418 ATGCTTAAGATTTGTTGTTGCAAGCATGCTGTTGCTATTATTGTGCCCTATAGATGCT 477
Db 574 TGGCGCCCAACTGACGCGCCCGACATAGACGTGCATTATGTGTGCATACAGATGCT 633
Qy 478 GAAGCATTGTAAGATAATATGCTCCCAATTTGCACTGTTGCTGCGCCAAACAATTTG 537
Db 634 CAGCAGCATTTAGCATATATCTCTGAATCATATGCAATCCGTTTGTATGAAACAGCAAT 693
Qy 538 GACTATGCAATTTTCAATGTTGAGAGCAAGTGGCAATCAGACGTTTAAATCGCGGAACTA 597
Db 694 GAATATGAAATTTTCAATCGTACGACAAAGAAACAAGAACTTCAACCGCCAAAGCTG 753
Qy 598 ATGAAGTTGATATAGAGCTAGCATCAGCCCTTACCCA-----TGGAGTGGCTTATC 651
Db 754 ATGAAGTTGGCTTGTGGAGAGCCAGAAAGTGTGGCGAAAGCTGGGAGTGCTTCTGAG 813
Qy 652 TTTTCATGATGCTGATTTATGCTGCCCGGAAGTGAACCGTAACTGTACAGTGTCCAAATTCAA 711
Db 814 TTTCACAGCATGACATGCTTGCTGCTCTCGACACACGAGAACCTGTATCTTCGCCACAGACAG 873
Qy 712 CCAAGTCATATAGATGCTGACGATCGATTAATTTCAATTAATTAATTTCAATATTTGGCGGATC 771
Db 874 CCGCGACATATGTCGGCTTCCATCGATTAATTAATTTCAATTAATTAATTTCAATTAATTTGG 933
Qy 772 TTGCGCGGAATCAGTGCATTAACAAAAGTCACTGAAAGAAATCAATGATTTTCGAT 831
Db 934 TTGCGCGGAGTCTTCCGATGACCTTTGAAAGTTCACAGAGGTCAACGGCTTCTCGAAT 993
Qy 832 GATTTTGGGGTTGGGGCGGAGAGACGATTTGGCGACGGAACATCGATGCTGGA 891
Db 994 AAGTACTGGGGTTGGGGCGGAGAGATGATGACATCTCATATAGTTAAAGAAATTAAT 1053
Qy 892 CTGAAAGTTTCAAGATATCCGACACAATTTGACGATTAATTAATTAATTAATTAATTAAT 951
Db 1054 TATCATATAGCAAGATACAAAATGTCTATAGCGCGGTACGCTATGTTAGATCATAGAAA 1113
Qy 952 GAAGCAGCAATTC 965
Db 1114 TCTACACTTAATTC 1127

RESULT 7
AR506512 1189 bp DNA linear PAT 22-SEP-2004
LOCUS AR506512
DEFINITION Sequence 11472 from patent US 6703491.
ACCESSION AR506512
VERSION AR506512.1 GI:52441987
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
 1 (bases 1 to 1189)
 Homburger, S.A., Ebens, A.J. Jr., Erickson, C.S., Francis-Lang, H.L.,
 Margolis, J.S., Reddy, B.P., Ruddy, D.A. and Buchman, A.R.
 Drosophila sequences
 Patent: US 6703491-A 11472 09-MAR-2004;
 Exelixis, Inc.; South San Francisco, CA
 Location/Qualifiers
 source 1. 1189
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN
 Query Match 16.6%; Score 191; DB 6; Length 1189;
 Best Local Similarity 58.6%; Pred. No. 8.9e-47;
 Matches 332; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

QY GCACGCCGGTGGACATGGAATGCTTAAGATTGTGTTCAGAGCATTCGTTGCTATTAT 458
 DB GGGCCCTGGTGGCGCTTCGAGCTGGAACCTGCAATGCCAGCATCAGCTGCTATTGT 567
 QY TGTGCTTATAGATCGTGAAGCACATTGGAATATGCTCCACATTTGCACTCGTT 518
 DB TGTGCTTATAGATCGTGAAGCACATTGGAATATGCTCCACATTTGCACTCGTT 518
 QY GCTGCGCAACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 578
 DB TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
 QY GTTTATCGCGGGAACCTAATGACGTTGATGATGATGATGATGATGATGATGATGAT 638
 DB CTTCATCGCGGCTGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 747
 QY GCAAGCTTCATCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
 DB GCAAGCTTCATCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
 QY GGATTTGTTTATATTCACAGATGATGATGATGATGATGATGATGATGATGATGAT 807
 DB GGATTTGTTTATATTCACAGATGATGATGATGATGATGATGATGATGATGATGAT 807
 QY GTGTCATTCACACACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 758
 DB CTGTTCACGTCACGCGCGGACATGATGATGATGATGATGATGATGATGATGATGAT 867
 QY ATATTCGGGATCTTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 818
 DB TATATTCGGGATCTTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 927
 QY TGGATTTTGAATGATTTTGGGGGTTGGGGGCGAGAGCAGCAGATTTGGCAGAGAAC 878
 DB TGGGTTCTCAAACTGCTTCTTTGGGCGGCGGCGAGATGACATGCTCCAGAGGTT 987
 QY ATGATGCTGAGCTGAAGAGTTTCAAGATATCCGACACAAATTCGACGATATTAATAT 938
 DB GAAGCAGCGCAACCTATTCATATCAAGATATCCGATGACATATGACCGCTACAGATGCT 1047
 QY TAAGCACTCGACGAGCGACGATCC 965
 DB 1048 GAAGCATCGAGAGGAAAGGCCATCC 1074

RESULT 8
 AY095531 1879 bp mRNA linear INV 17-APR-2002
 LOCUS Drosophila melanogaster RES5531 full insert cDNA.
 DEFINITION AY095531
 ACCESSION AY095531.1 GI:20177118
 VERSION
 KEYWORDS FLI_CDNA.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 1879)
 Stapleton, M., Brokstein, P., Hong, L., Asgayani, A., Carlson, J.,
 Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Friese, E.,

George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G.,
 Miranda, A., Mungall, C.J., Nuno, J., Pacle, J., Paragas, V., Park, S.,
 Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S.B., Rubin, G.M.
 and Celisner, S.
 Direct Submission
 Submitted (11-APR-2002) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720
 This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy, presence of a polyA tail and contiguity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcription of unspliced precursor RNAs, and
 reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our web site
 (<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES
 source
 1. 1879
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /strain="Y: cn bw sp"
 /db_xref="taxon:7227"
 /map="50B9-50B9"
 1. 1879
 /gene="BCDNA:GH13356"
 /note="Alignment with genomic scaffold AE003815"
 /db_xref="FLYBASE:FBgn0027538"
 133..1344
 /gene="BCDNA:GH13356"
 /note="Longest ORF"
 /codon_start=1
 /product="RE5531P"
 /protein_id="AAM12262.1"
 /db_xref="GI:20177119"
 /db_xref="FLYBASE:FBgn0027538"
 /translation="MYLFTKANLIRFLAGALCLLVNFGFRSDGSGATSLSKLSIR
 RVHRYAHYVGNASDAGSSEASRLPASPLASDREDOELNGPNTIRTVATAN
 FTSIPDLTRFLGKFLPPROKSTSLANCTDPPRDGAPTPNTIRSLDVEA
 ELGSLIRRGGAFFEPENCNOHVAIVPFRRYVHLILFLNHPFLMKORLAVIF
 VEQNGKRPFRNAAAMNIGYLEAKLKYQWDCIRPHVDLLPLDDRLNKPQPRMVS
 AIDTNFRPLPRSTIFGVSAATREHFOVNPSPNSFRWGGSDDMSRNLHNPANFIS
 RYPVNIARVKNLKHQKEKANPKRYENLQNGMSKLEBDGINSIKYSIYSIKOPPTTWY
 LAELKNSERKS"

ORIGIN
 Query Match 16.6%; Score 191; DB 2; Length 1879;
 Best Local Similarity 58.6%; Pred. No. 9.4e-47;
 Matches 332; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

QY GCACGCCGGTGGACATGGAATGCTTAAGATTGTGTTCAGAGCATTCGTTGCTATTAT 458
 DB GGGCCCTGGTGGCGCTTCGAGCTGGAACCTGCAATGCCAGCATCAGCTGCTATTGT 689
 QY TGTGCTTATAGATCGTGAAGCACATTGGAATATGCTCCACATTTGCACTCGTT 518
 DB TGTGCTTATAGATCGTGAAGCACATTGGAATATGCTCCACATTTGCACTCGTT 749
 QY GCTGCGCAACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 578
 DB TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 809
 QY GTTTATCGCGGGAACCTAATGACGTTGATGATGATGATGATGATGATGATGATGATGAT 638

[illegible]

RESULT 9			
AB036858	2796 bp	mRNA	INV 15-JUL-2000
LOCUS			
DEFINITION	AB036858		
	Clona intestinalis Cl-b4Gal-T mRNA for beta 4		
	galactosyltransferase, complete cds.		

VERSION	AB036858.1	GI:9229931
KEYWORDS	Ci-b4Gal-T; beta 4 galactosyltransferase.	
SOURCE	Ciona intestinalis	
ORGANISM	Ciona intestinalis	

REFERENCE AUTHORS	TITLE
Phlebotominae; Cionidae; Ciona. 1 (sites)	
Hotta, K., Takahashi, H., Asakura, T., Satoh, B., Takatori, N., Satou, Y. and Satoh, N.	Characterization of Brachyury downstream notochord genes in the Ciona intestinalis embryo

**JOURNAL
REFERENCE
AUTHORS
TITLE**
JOURNAL

Ciona intestinalis embryo
Unpublished
2 (bases 1 to 2796)
Hotta, K., Takahashi, H. and Satoh, N.
Direct Submission
Submitted (13-JAN-2000) Kohji Hotta, Graduate School of Science,
Kyoto University, Department of Zoology, Sakyo-ku, Kyoto, Kyoto,
606-8501, Japan [E-mail: khoti@sci.kyoto-u.ac.jp]
Tel: +81-75-753-4095, Fax: +81-75-705-1113

FEATURES	
source	location/Qualifiers
gene	1..2796
	/organism="Ciona intestinalis"
	/mol_type="mRNA"
	/db_xref="taxon:7719"
	/cell_type="notochord"
	/note="This gene is expressed in the notochord cells"
	1..2796
CDS	/gene="Ci-bdGal-T"
	37..1278

```

/translation="WSLRHQVKGIVKRTNYGLRAKMLTCLILFALEIVIVFRIH
DVIILSIHLNITKLMIRGLETFNCEAVGSHNSLKOSRLANGENASPLPCPEMREH
LLKQELGNEDEVSLLDVKNYTLVEPEGREFKPAQCVPRHKVAILVPHRRHRLR
FLKA IHPVMQROADYGVFVHVSQGTGFENFKKALITGTEALKDDYDCLFPHVDI
LAEDDRNLT YRCADVPRHLSVGI DKNDYDLPDALFGVIAITKQDFAQVNGVSNENYC

```

Query Match	16.2%	Score 186.8;	DB 2;	Length 2796;
Best Local Similarity	53.6%	Pred. No. 1.9e-45;		
Matches 389;	Conservative 0;	Mismatches 337;	Indels 0;	Gaps 0;

ORIGIN

```

WGAEDDDMYRRIIHSICLIGERADQYDVARYMAAHPSPKSNRVNYPRTLLVGAEROR
HDEGNINHHYSLVEKIQQLPLTYNISADYGRPLNPKYVGFOTGIDIVYSISLLMLTLM
TCCTCKVSRLLPHVLCPRIIP"

```

QY	354	CGAAGACCCCGCACTTCAAAACTCGAGAAAATCTATCCGGACACGCGACCGGTGGACA	413
	375	CGAGACGTTAGCTTGGACCATCTAGTAAAAAACTACACGTTGGTGAACCTGGTGGGCG	434
QY	414	TGGAATGCCCTTAAGGATTGTGTGGAAAGGATCGTGTTCATTATTTTGGCCCTTAAGA	473
Db	435	GTTTCAAAACCGGCGCAATGTGTCCCTCGTCACAAGGTGGCAATTATCGTCCGCACTTAA	494
QY	474	TCTGGAAGCAATTTGGAATTAATGCTCCCAATTTTGCATCTCGTGTCCGCAACACA	533
Db	495	CCGTGAGAGACACTACGGCAGTTTCTTAAAAAGCTATACCCCTGTATTGACAGCGCAGCA	554
QY	534	ATTGGACTATGCAATTTTCAATTGTGAGCAAGTGGCGCAATTCAGCGTTTAATCGCGGAA	593
Db	555	AGCGCACTATGATGTTTTCGTGTGTTCAACCAATCCGGCACCGGCACATTTTAAACAAGCAAA	614
QY	594	ACTAATGACGTTGGATACGACGTAAGCATCACGCGCTTACCCATGGGAGTCTCATCTT	653
Db	615	GCTGTTAAATATCGGCTACACTGAAGGCTGGAACAGATGATTATGATTTGCTTCACTTT	674
QY	654	TCATGATGTCGATTACTGCCCCGGAAGATGACCGTAACTGTACACGTCCAAATTCAAC	713
Db	675	TCAAGATGTGACCTGCTCGCTGAGGACGACCGCAATCTGTACCGATGTGCGCATGTAC	734
QY	714	ACGTCATATGAGTGTAGCGATTCGATTAATTCATTTATTAACCTTCATATTCCGGGACTTT	773
Db	735	CAGACATTTATCTGTGTGTATCGCAAAATGGGATTTACCAATGGCATATGACGCACTGTT	794
QY	774	CGGCGGATCATGTCACCTAAACAAAGATTCACCTGAAGAAAATCAATGGATTTTTCGAATGA	833
Db	795	CGGTGGGATTTTTCATATGACCAAAAGACCAATTTGCCAGGTCAACGGGTATTTGAAACGA	854
QY	834	TTTTTGGGGTTGGGGCGAGAGAGACGACGATTTGGCGACGAAACATCGATGCTGCACT	893
Db	855	ATTACTGGGGATGGGGTGCAGAAAGATGACGACATGATGTGAGGATTCCTGCATACCTGTCT	914
QY	894	GAAAGTTTCAGATATTCGCACACAAATTTGCACGATATTAATGATTAAGCACTGAGCGGA	953
Db	915	TGCGTTTAAAGAACACAAATATGACGTTGACGGTATCGAATGGCTTACCAACCTTTCGGA	974
QY	954	AGCGACGAATCCAGTTAATAATACCGCTACCAAAATATGGCGCAAAAGAACGCCCGATG	1013
Db	975	CAAAAGCAACCGAATTAATTCATATTCGTTACCACTTACTGTTGGCGCGCGGAGCGACA	1034
QY	1014	GACACGTGACGGCTTAAGCAATCTGAAAGTAAAGCTCGTAATCTCGAATTTGAAGCCTCT	1073
Db	1035	GAGCGATGATGATTAATAACAACCTTCAATTACAGCTGTGTAGAAAAAACACAGCTACCATT	1094
QY	1074	CTACAC 1079	
Db	1095	ATACAC 1100	

RESULT	10				
AKI16794					
LOCUS	AKI16794	2082 bp	mRNA	linear	INV 30-NOV-2002
DEFINITION	Ciona intestinalis cDNA,	clone:c1eg305d14,	full	insect	sequence.
ACCESSION	AKI16794				
VERSION	AKI16794.1	GI:23589473			
KEYWORDS	FLI-CDNA.				
SOURCE	Ciona intestinalis				
ORGANISM	Ciona intestinalis				
	Euryarchaeota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;				
	Phlebobranchia; Clonidae; Ciona.				

REFERENCE	AUTHORS	TITLE	JOURNAL	PUBMED	REFERENCE
1	Satou, Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T., Sasaki, A., Hamaguchi, M., Awazu, S., Yagi, K., Sasaoka, Y., Nakayama, A., Ishikawa, H., Inaba, K. and Satoh, N.	A CDNA resource from the basal chordate <i>Ciona intestinalis</i>	Genesis 33 (4), 153-154 (2002)	1203911	
2	(bases 1 to 2082)				
	Satou, Y. and Satoh, N.	Direct Submission			
	Submitted (04-OCT-2002)	Nori Satoh, Kyoto University, Department of Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan			
	Ex-mail: satoch@esdian.zool.kyoto-u.ac.jp, Tel: 81-75-753-4095, Fax: 81-75-705-1113				
	<i>Ciona intestinalis</i> CDNA Project (URL: http://ghost.zool.kyoto-u.ac.jp/indexr1.html).				
FEATURES	source	Location/Qualifiers			
	1..2082	/organism="Ciona intestinalis"			
		/mol_type="mRNA"			
		/db_xref="taxon:7719"			
		/clone="c1eg05d14"			
ORIGIN					
Query Match	15.7%;	Score 181.4;	DB 2;	Length 2082;	
Best Local Similarity	54.2%;	Pred. No. 8.6e-44;			
Matches 368;	Conservative 0;	Mismatches 311;	Indels 0;	Gaps 0;	
QY	401	ACGCCGGTGACATGGAATGCCCTAAGATGCTGTGCAGGCACTGCTGCTATATTG	460		
DB	853	AAGATGGGGGACCTACCAAGCCCCCTTACTGTAAAGCCTTATGTAAAGTCCCATGTATA	912		
OY	461	TGCCTTATAGATCGTGAAGCACATTTGGAATAATAGCTCCACATTTGGCACTGTTGC	520		
DB	913	TTCCTTATCGGACAGAGAAACATTTGCGGTATTTCTTGAGTACATGATCCCAACC	972		
OY	521	TCGCCAAACAAATTTGGACTATGCAATTTTCATTTGTGAGACCAATGGCGAATCAGAGT	580		
DB	973	TACAAAGCACAACCTGATTACGCATCTATGTTGTTATACAGGCTGGTACTGGTAAT	1032		
OY	581	TTATTCGGGGAAACTATGAACGTGTGATACAGGTAGCATCAAGCCTCTATCCATGTCGC	640		
DB	1033	TTAATTCGGCCACACTGATGATGTGTGTATGCTGACATCAATCAAGATCATGATTTTC	1092		
OY	641	AGTGCCTCATCTTTCATGATGTGATTTTACAGCCGGAAGATGACCGTAACTGTACAGT	700		
DB	1093	AATGTTTTCCTTCCATGATGTGACTTGTGTCTTGAGAAATGACAAGTATATATCTTT	1152		
OY	701	GTCGAATTCACAACGTCATATGATGATGAGTGAAGCATGATTAATTCATTTAACTTCAT	760		
DB	1153	GCCCCCAAGTCCTCGCCACCTGTACAGCTGGGGGTGAGACAGATTTAATTAACCACTGCCCT	1212		
OY	761	ATTGCGGAGCTTGTGGCGGAGTCACTGATCACTAAACAAAGATCACTGAAAGAAATCATG	820		
DB	1213	ACTCTGCCATATTTGTGTGCGCTGACTGCTGACCAAGAAACAAATTCACAAAAGTGAAGC	1272		
OY	821	GATTTTCGAATGATTTTGTGGGTTTGGGCGGAGAGACGACGATTTGGCGACGAACAT	880		
DB	1273	GTTACTCTAATTCATTTCTGGGGTTGGGGTGGAGAAAGACAGATATGTTTAAACGCTGCA	1332		
OY	881	CGATGGCTGACCTGAAAGTTTCAAGATATCCGACACAAATTGCAACGATATTAATGATTA	940		
DB	1333	AGTTTTCGGGATGATATTTATTTGTTATCCCTATGAGATATCTCAAGGATATTAAGATTA	1392		
OY	941	AGCACTGCAGCGAAGCGACGATCAGTTAATTAATTTGCGCTACAAATTAATGGCCAA	1000		
DB	1393	CTCATCAAGAGAGAAAGGAAACGAACCAAAACCTTGACCAAAATACGACAGA	1453		
OY	1001	CGAAGCGCGATGACACGTACGGCCTTAAGCATCTGAAGTATAGCTGTTAAATCTGG	1060		
DB	1453	CTTAAAGACACATGGCAACGATGGTCTTAAACACACTTGAATATACAGGTCTGATCAAGC	1512		
OY	1061	AATTGAAGCCTCTCTACAC	1079		

Db	1513	AGAGAAGCAAGCTTATAC	1531
RESULT 11			
LOCUS	BC077601		
DEFINITION	Xenopus laevis UDP-Gal:betaGalCNC beta 1,4-galactosyltransferase, polypeptide 3, mRNA (CDNA clone MGC:84507 IMAGE:5077911), complete cds.		
ACCESSION	BC077601	2265 bp	MRNA linear VRT 12-OCT-2004
VERSION	BC077601.1	GI:50417572	
KEYWORDS	MGC.		
SOURCE	Xenopus laevis (African clawed frog)		
ORGANISM	Xenopus laevis		
REFERENCE	Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Piproidea; Pipidae; Xenopodinae; Xenopus; Xenopus.		
AUTHORS	1 (bases 1 to 2265) Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.		
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative		
JOURNAL	Dev. Dyn. 225 (4), 384-391 (2002)		
PUBMED	12454917		
REFERENCE	2 (bases 1 to 2265) Strausberg,R.L., Peingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Stennmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Rubin,G., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schaeetz,T.E., Brownstein,M.J., Uedini,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mulrany,S.J., Bosak,S.A., McBean,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Huylk,S.W., Villalón,D.C., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Halton,E., Kettman,M., Madan,A., Young,A.C., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Madan,A., Rodriguez,S., Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,U., Myers,R.M., Buttefield,Y.S., Krzywinski,M.I., Skalske,U., Smalls,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12477932		
REFERENCE	3 (bases 1 to 2265) Klein,S. and Gerhardt,D.S.		
AUTHORS	Direct Submission		
TITLE	Submitted (19-JUL-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA		
JOURNAL	NIH-MGC Project		
PUBMED	Contact: XGC help desk		
REFERENCE	Email: cgapps-remail.nih.gov		
AUTHORS	Tissue Procurement: Martha Rabbett, Steven L. Klein, Ph.D.		
TITLE	cDNA Library Preparation: Life Technologies, Inc.		
JOURNAL	cDNA Sequencing by: The I.M.A.G.E. Consortium (ULNL)		
PUBMED	BC Cancer Agency, Vancouver, BC, Canada		
REFERENCE	info@bcgsc.bc.ca		
AUTHORS	Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Buttefield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Porin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Pribbu, Parvaneh Saeedi, JR Santos, Angelique Scherck, Ursula Skalska, Duane Smalls, Jeff Stolt, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.		

BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Anara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lisa Prabhu, Parvaneh Saeedi, JR Santos, Angeliq Schercher, Ursula Skaleka, Duane Smalus, Jeff Stott, Miranda Teali, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRAX Plate: 156 Row: 1 Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source

1. 2681
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:81163 IMAGE:6318052"
/tissue_type="Ovary, Xenopus"
/clone_1ib="NICHD_XGC_Ov1"
/lab_host="DH10B"
/note="Vector: PCMV-SPT06"
1. 2681
/gene="MGC81163"
/db_xref="GeneID:414540"
77. 1573
/gene="MGC81163"
/codon_start=1
/product="MGC81163 protein"
/protein_id="AAH68719.1"
/db_xref="GI:46250086"
/db_xref="GeneID:414540"
/translation="MMRRVLEPCTLAMLIGFOPAFMYFSGRLSLFIRGHSIEP
SDYSRTHDVYNSLSVNSMPTSTAEIILLPRCPSPRYLVGPIITTFSPSPSLK
VDDKPLVTKGEYKPCNEAHRITAVIIPHNRETHLHLYLHPLOOOLHRI
YIHQNGSTFRKALINVEKALRDEMDOLFHDVLIPEDFNLVPCWSPK
ASVANKFSYVLPYMFYFGVSLTPDQYMKNGFPEYWGSGEDDIDATVRLAG
KITRPSVAGHKMGHGDHNEENPHFDLIRQRMKTDGMSLKYLABELE
HLYTVTVDIGVDPRAKKLRGRGPPPIRRYVVKSTSTVSQSIQNDLTLVSE
INGVERSAKKRPVVAPESTTTTBKSLPIASAPRRRVVDKQIRYIRSTEDNNKAQV
KEVQADVAKRGARSSKLERQVDSFRAPQDVGKSLPQSLSRAPV"

gene

CDS

ORIGIN
Query Match 14.8%; Score 170.2; DB 5; Length 2681;
Best Local Similarity 51.4%; Pred. No. 2.5e-40;
Matches 419; Conservative 0; Mismatches 393; Indels 3; Gaps 1;
QY 286 GACAAACAGACTCTTCAATTCGTATCAACACCTCCCACTCGTGGACCCATCCGT 345
DB 293 GACATCTGCTCTTGAATGCCAGAGAGGTACCTTATTTGGTGGCCCAATTACCGTC 352
QY 346 GATTTCCTGAG 405
DB 353 ACCTTAGCCAAACCCCTCCCTTAAAGAGTTCAAGATTAAGAACCCCTGTGATPAAA 412
QY 406 GGTGACATGAGATGCTTAAGATGTGTGCAAGAGCATCGTGTCTATTTATTTGGCCC 465
DB 413 GGTGGGAGTACAG 472
QY 466 TATAGATCGTGAACACATTTGAGATTAATGCTCCACATTTGCACTGTTGCTCGCC 525
DB 473 CACCGGAG 532
QY 526 AAACAAATTTGAGATGAGATTTTCAATTTGAGAGCAAGTGGCAATGAGAGCTTTAAT 585
DB 533 CGCAGAGAGTCCACATGAGATCTACATCATCAGGCGGGTAACTCCATTTAAT 592
QY 586 CGCGGAGAACTAATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 645

DB 593 CGGCGCAAGCTGCTGAATGTTGGGTGAAGAGCTCGCGGAGCAGAGACTGGAGCTGT 652
QY 646 TTCAATCTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 702
DB 653 CTGTTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 712
QY 703 CCAATTAACACAGCTCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 762
DB 713 CCTGAGAGCCCAACAG 772
QY 763 TCGGAGATCTTCGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 822
DB 773 CCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832
QY 823 TTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 882
DB 833 TTTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892
QY 883 ATGCTGAGACTGAAGATTTCAAGATATCCGACACAAATTTGACGATATTAATGATTAAG 942
DB 893 CTGGCCGATGATGAATATACCCGCCGCTCGTGGCGGCGCACTATTAAGATGATGATGAT 952
QY 943 CACTGAG 1002
DB 953 CACAAGGAG 1012
QY 1003 AAGCCCGATGAGACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1062
DB 1013 CAGAAATATGATGAG 1072
QY 1063 TTGAAGCCTCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097
DB 1073 CTGGAACATCTTACACCAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1107

RESULT 13
AY117537
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1338
/organism="Cricetulus griseus"
/mol_type="mRNA"
/db_xref="taxon:10029"
/tissue_type="ovary"
52. 1239
/note="beta4galT-3"
/codon_start=1
/product="beta-1,4-galactosyltransferase 3"
/protein_id="AAH77196.1"
/db_xref="GI:30313426"
AY117537 1338 bp mRNA linear ROD 01-MAY-2003
Cricetulus griseus beta-1,4-galactosyltransferase 3 mRNA, complete
cde.
AY117537
AY117537.1 GI:30313425
Cricetulus griseus (Chinese hamster)
Cricetulus griseus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Cricetidae; Cricetinae; Cricetulus.
Lee, J., Park, S.-H., Sundaram, S., Shaper, N.L., Raju, T.S. and
Stanley, P.
Consequences of the lower expression level of six
beta-1,4-galactosyltransferase transcripts in Lec19 CHO
glycosylation mutants
Unpublished
2 (bases 1 to 1338)
Lee, J., Park, S.-H. and Stanley, P.
Direct Submission
Submitted (29-MAY-2002) Cell biology, Albert Einstein College of
Medicine, 1300 Morris Park Ave., Bronx, NY 10461, USA
location/Qualifiers
1. 1338
/organism="Cricetulus griseus"
/mol_type="mRNA"
/db_xref="taxon:10029"
/tissue_type="ovary"
52. 1239
/note="beta4galT-3"
/codon_start=1
/product="beta-1,4-galactosyltransferase 3"
/protein_id="AAH77196.1"
/db_xref="GI:30313426"

QY	281	CCCCCGAACACAACTCTTCAATTTCGTATATCAACAACCTCCGCACTGCTGGACCA	340
Db	426	TCCCAAGCTCTCAAAGGCTGCACTACGTGCCCAACGATCTCTCTTATAGGGTCCCA	485
QY	341	TCCGTGTAATTCCT---CGATGAGCCCGACTTCGAAAACCTCTGAGAAAAATCATCCGACA	397
Db	486	TATCCGTGTCCTTTAGCCCAAGTGCATCGCTGGCAGAGATGTGTGAGAGGAATCCCGGG	545
QY	398	CGCACGCCGGTGGACATGGAATTCCTTAAGATTTGTGTGGCAAAGCATCTGTGTCTATTA	457
Db	546	TAGAACCAAGGGGGCGATACCGCCCTCGAAGGTGTAGACCCCGGTCCCGACACCACTCA	605
QY	458	TTGAGCCCTTATAGAGATCGTGAAGACAACTTTGAGAAATPATCTGCACAAATTGGACTCGT	517
Db	606	TTTGTCCTCCCAACCGTGCCCGGGAGACACACTGCGCTGCTCTCTTACCACTGCAACCCGT	665
QY	518	TGCTTCGCCMAACAAATYGAAGTATGCAATTTTCATTTGGAGCAAGTGCGAATCAGA	577
Db	666	TCTGACAGCGCAGAGCTGCTTACAGGCATTTATGTATCACAACAGGCTGGAATGAAA	725
QY	578	CGTTTAATCCGGGGAACTTAATGAACGTGTGGATACAGCTGTGCATCAGCCTTAACCAT	637
Db	726	CATTTAACAGGSCAAAGCTGCTGATGTGTGGGGTGGCGGAGGCCCTGGAGATGAAGAT	785
QY	638	GGCAGTGCCTTATCTTTGATGATGTCATTTACTGCCCCGAGATAGCCGTAACTGTACA	697
Db	786	GGAGCTGTTGTTTCTGATGATGTGTGATCTCTGCGCAGAGAAAGCAACCACTGTATG	845
QY	698	CGTG---TCCAATTCACCAACGTCATATGAGTGTAGCGATGATTAATTCATTTATTAAC	754
Db	846	TGTGTGATCCCCCGAGGACCTCGGCACTGTGCTGTGGCCATTAACAAATTGGAATAGCC	905
QY	755	TTCCATATTGCGGCATCTTTCGCGCGGAATCAAGTGAAGTAAACAAAGATCACTGAAGAAA	814
Db	906	TCCCGTATCCCCAGTACTTTTGGCGGGGTCTCAGCGCTCACTTCGACCAAGTAACTGAAAG	965
QY	815	TCAATGGAATTTTCGATGATTTTGGGGTTGGGGCGGAGAGACGACGATTTGGCGAGA	874
Db	966	TGAATGGCTTCCCCAATGATATACGTGGGCTGGGGGTGGTGAAGATGACGACATTCGTACA	1025
QY	875	GAACTCATGTGGCTGGACTGAAAGTTTCAAGATATCCGACCAAAATTCACAGATATATAA	934
Db	1026	GGGTTTCGCTTGGCTGGAGTGAAGATATTCGTCCCCCAACATCCGTGGGACACATATAGA	1085
QY	935	TGATTAAACACTCTCGACGGAAGGACGAGATCCAGTTATTAATATGCGCGTACAAAATTAATG	994
Db	1086	TGTGTAAAGCACCGAGGAATTAAGGGCAACGAGAAAAATCCCCACAGTTTGACTCTCTGG	1145
QY	995	GCCAAACGAAGCGCCGATGACACGCTGACGGCTTAAGCAATCTGAAGTATTAAGCTCGTAA	1054
Db	1146	TCCGTAAACCAAGAAATTCGTGACTCAAGATGGAATGAATCTCACTGACATTAACAGTTGCTGT	1205
QY	1055	ATCTGGAATTTGAAGCTCTCTTACAC	1079
Db	1206	CTCGAGAGCTGGCCCTCTCTATAC	1230

Search completed: January 8, 2006, 21:00:07
Job time : 6023 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, (using frame_plus.p2n model)

Run on: January 9, 2006, 00:31:56 ; Search time 6343 Seconds
(without alignments)
3432.294 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSILVLCV.....VLLKDKRRRLRDPPTCF 383

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5883141 segs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlh
-MODE=frame.p2n.model -DEV-xlh
-DB=gen2.1/USPTO.spool/US10661430/runat_06012006_123134_6236/app_query.fasta.1.519
-DB=Genemb1 -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10661430.GCNC.1.1.4939 @runat_06012006_123134_6236 -NCPU=6 -ICPU=3
-NO MAP -LARGEOUTER -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Genemb1:.*
1: gb_ba:.*
2: gb_in:.*
3: gb_env:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pr:.*
9: gb_ro:.*
10: gb_scs:.*
11: gb_sy:.*
12: gb_un:.*
13: gb_vl:.*
14: gb_hlg:.*
15: gb_pl:.*

Pred. No. is the number of results predicted by change to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	1152	2	AY130767 Caenorhab
2	2031	100.0	1152	2	AY533306 Caenorhab
3	1275	62.8	63747	2	AC025727 Caenorhab

4	775.5	38.2	1269	2	AY601103	AY601103 Trichoplui
5	741	36.5	2265	5	BC077601	BC077601 Xenopus t
6	740	36.4	2098	5	CR942597	CR942597 Xenopus t
7	736.5	36.3	2681	5	BC068719	BC068719 Xenopus t
8	731.5	36.0	2082	2	AK116794	AK116794 Clona Int
9	730	35.9	2626	5	BC041742	BC041742 Xenopus t
10	729	35.9	2340	5	BC075452	BC075452 Xenopus t
11	725	35.7	1952	9	AF142671	AF142671 Mus muscu
12	725	35.7	1962	9	BC013619	BC013619 Mus muscu
13	721	35.5	1927	4	BT020682	BT020682 Bos tauru
14	720	35.5	1329	5	BC087436	BC087436 Xenopus t
15	720	35.5	1921	4	BT020916	BT020916 Bos tauru
16	720	35.5	1949	4	BT020868	BT020868 Bos tauru
17	717	35.3	1182	8	HSUDPAL	Y12509 Homo sapien
18	717	35.3	1193	8	AB024435	AB024435 Homo sapi
19	717	35.3	1338	9	AY117537	AY117537 Cricetulu
20	717	35.3	1912	8	BC009985	BC009985 Homo sapi
21	717	35.3	1915	8	BC000276	BC000276 Homo sapi
22	717	35.3	1934	8	BC006099	BC006099 Homo sapi
23	717	35.3	1938	6	BD206381	BD206381 Human nuc
24	717	35.3	1938	6	AR400696	AR400696 Sequence
25	717	35.3	1938	6	AX013200	AX013200 Sequence
26	716	35.3	2007	9	BC061812	BC061812 Rattus no
27	714	35.2	1885	8	AK092493	AK092493 Homo sapi
28	709	34.9	1839	6	CO590796	CO590796 Sequence
29	709	34.9	1879	2	AY095531	AY095531 Drosophil
30	709	34.9	2269	2	AF132158	AF132158 Drosophil
31	706	34.8	1920	8	AF038661	AF038661 Homo sapi
32	704	34.7	1189	6	AR506512	AR506512 Sequence
33	694	34.2	1912	8	AF020921	AF020921 Homo sapi
34	684.5	33.7	1954	5	CR760293	CR760293 Xenopus t
35	684	33.7	1853	6	BD191505	BD191505 Secreted
36	680.5	33.5	2279	5	GGU19890	U19890 Gallus gall
37	675	33.2	4093	6	CO590795	CO590795 Sequence
38	675	33.2	35217	14	AC017131	AC017131 Drosophil
39	675	33.2	165177	2	AC007588	AC007588 Drosophil
40	675	33.2	180425	2	AC008229	AC008229 Drosophil
41	675	33.2	229685	2	AE003815	AE003815 Drosophil
42	671.5	33.1	1262	9	AF118896	AF118896 Cricetulu
43	670	33.0	2420	5	GGU19889	U19889 Gallus gall
44	668.5	32.9	2304	6	A38781	A38781 Sequence 7
45	668.5	32.9	2304	6	I49897	I49897 Sequence 7

ALIGNMENTS

RESULT 1
AY130767 1152 bp mRNA linear INV 17-SEP-2002
LOCUS
DEFINITION
{beta}1,4-N-acetylgalactosaminyltransferase mRNA, complete cds.
ACCESSION
AY130767
VERSION
AY130767.1 GI:22415754
KEYWORDS
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae;
Rhabditidae; Rhabditidae; Pelodidae; Caenorhabditis.
REFERENCE
AUTHORS
TITLE
Kawar, Z.S., Van Die, I. and Cummings, R.D.
Molecular Cloning and Enzymatic Characterization of a
UDP-GalNAc:GlcNAc-beta-1,4-N-Acetylglucosaminyltransferase
from Caenorhabditis elegans
J. Biol. Chem. 277 (38), 34924-34932 (2002)

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
Kawar, Z.S., Van Die, I. and Cummings, R.D.
Direct Submission
Submitted (10-JUN-2002) Biochemistry and Molecular Biology,
University of Oklahoma Health Sciences Center, 975 NE 10th St. BRC
Rm. 417, Oklahoma City, OK 73104, USA
FEATURES
source
1. 1152
location/Qualifiers

CDs
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
1..1152
/note="Ce(beta)4GALNAC1; similar to Caenorhabditis elegans
hypothetical protein Y73E7A.7"
/product="UDP-GalNAc:GlcNAc(beta)-R
{beta}1,4-N-acetylglucosaminyltransferase"
/protein_id="AA095168.1"
/db_xref="GI:22415755"

ORIGIN
Alignment Scores:
Pred. No.: 2 84e-179 Length: 1152
Score: 2031.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-10-661-430-1 (1-383) x AY130767 (1-1152)
QY 1 MetAlaPheArgHisLeuAlaValAlaArgLeuSerLeuValLeuCySaIaVal 20
Db 1 ATGGCTTTTCGTACTTTGGCAGTGCAGACTCAAGTCGTTGCTCGTACTTTGGCGCTT 60
QY 21 LeuLeuValHisAlaMetIleTyrLysIleProSerLeuTyrGluAsnLeuThrIle 40
Db 61 CTTCATTTAGTCAATGCAATGATTATTAAGATTCATCGCTTTCAGAACTTACTATC 120
QY 41 G1SerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla 60
Db 121 GGCTCTTCGACCTTATTTGCCAGCTGAGCGCAATGAGGAGTCTCTGGGAATACGGCT 180
QY 61 SerThrSerAspLeuAspThrTyrAsnSerThrPheSerProIleSerGluVal 80
Db 181 TCCACTTCGAGTACTTACTTGATGAGTGGAAATCCAGTTTTCACGATTTTCGAAGTT 240
QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
Db 241 AATCAGACTAGTTTATGAGGACATTCCTCAATCCTGTTCCCGCAACACGACTCTT 300
QY 101 G1PheCysAsnGlnThrProProHisIleValGlyProIleArgValPheLeuAspGlu 120
Db 301 CAATTCTGTAATCGACACCTCCCACTCTCGACCCATCCGTGTAATCTTCGATGAG 360
QY 121 ProAspPheLeuThrLeuGluLysIleTyrProAspThrHisAlaGlyHisGlyMet 140
Db 361 CCCGACTTCAAACTCTCGAGAAATCTATCCGACACGCGCGGAGGACATGGAATG 420
QY 141 ProLysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGlu 160
Db 421 CCTTAGAGATTGTTGGACAGGCACTGTTGCTTATTTGTCCTTATAGAGATCGGAA 480
QY 161 AlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnGlnLeuAsp 180
Db 481 GCACATTGGAATAATGCTCACAAATTGCACTGTTGCTCGCAACCAACAAATTGAC 540
QY 181 TyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMet 200
Db 541 TATGCATTTTTCATTTGGAGCAAGTGGCGAATCAGAGCTTATTCGGGGAAACTAATG 600
QY 201 AsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAsp 220
Db 601 AACGTGGATACGACGTACATCAGCCTTACCACTGAGCTTCACTTTCATCATGAT 660

QY 221 ValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHis 240
Db 661 GTCGATTACTGCCGCAAGATGACCGTAACTGTAACAGCTGTCAATTCACACGCTCAT 720
QY 241 MetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGly 260
Db 721 ATGAGGTGAGCGCATCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
QY 261 IleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTyr 280
Db 781 ATCAGTGCCTCAACAAAGATCACTGAAGAAATCAATGATTTTCGAATGATTTTGG 840
QY 281 G1TyrG1 300
Db 841 GGTGGGGGGGAGAGGACGACGATTTGGGAGGAGAAATCATGCTGCTGAGCTGAAGTT 900
QY 301 SerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThr 320
Db 901 TCAGATATCCGACACAAATTCGACGATTAATTAATTAATTAATTAATTAATTAATTAAT 960
QY 321 AspProValAsnLysCysArgTyrLysIleMetG1G1ThrLysArgTyrThrArg 340
Db 961 AATCCAGTTAAATTAATGCGCTCAAAATTAATGAGGCAACGAGGCGGATGACACGT 1020
QY 341 AspG1LysSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThr 360
Db 1021 GACGCGCTTAAGCATTGAAGTAACTGTAAGCTGTAATCTGGAATTAAGCTCTCAACCT 1080
QY 361 ArgAlaValAlaAspLeuGluLysAspCysArgArgG1G1LeuLysArgAspPhePro 380
Db 1081 CGAGCGCTGCTGATTTGCTCGAAAAAGACTGCCGCGGAGCTGCGAAGGAGCTTCCA 1140
QY 381 ThrCysPhe 383
Db 1141 ACGTGTTT 1149
RESULT 2
AY533306 1152 bp mRNA linear INV 23-FEB-2004
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS
Griffiths,J.S., Huffman,D.L., Whitacre,J.L., Barrows,B.D.,
Maroquin,L.D., Muller,R., Brown,J.R., Henne,T., Bako,J.D. and
Aroian,R.V.
Resistance to a bacterial toxin is mediated by removal of a
conserved glycosylation pathway required for toxin-host
interactions
J. Biol. Chem. 278 (46), 45594-45602 (2003)
12944392
2 (bases 1 to 1152)
REFERENCE
AUTHORS
Griffiths,J.S., Huffman,D.L. and Aroian,R.V.
Direct Submission
Submitted (26-JAN-2004) Biology, Univ. California, San Diego, 9500
Gilman Dr, La Jolla, CA 92093, USA
location/Qualifiers
1..1152
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/db_xref="taxon:6239"
/chromosome="I"
/map="Y73E7A"
1..1152
/gene="bre-4"
1..1152
/gene="bre-4"
CDs
/gene="bre-4"

/note="glycosyltransferase; required for susceptibility to
Bacillus thuringiensis Cry toxins"
/codon_start=1
/product="BRE-4"
/protein_id="AA521308.1"
/db_xref="GI:4260122"
/translation="MAFRHLAVARLKSLLVLCVALLVHMYTKIPSLYENLTIGSST
LIADVDAMEAVLGNTPASTDDLDITWNSYFSPISSEVNTSPMEDIRPILFPNQTLQF
CNOETPHLVGPIRVLPDEDPFKLEKIVDTAGGGHMPKDCVARRHVAIVYVRE
ALRIIMLHSLAKKQGLDVAIPVEOVANOTFRGKLMYGVASVLSYVWOCETP
HDVLLPEDDRLLYTCPIOPRMSVAIDENYKLPISA.FGSI.SALT.KDHLKINGS
NDFWKGESDDLATRTSMAGKRSRYTPQIARYKMTKSTEAIPVNCCKRIMQGT
KRRWTRDGLSNLKYKLVNLELPLVYTRAVVDLEKDCRELARDPEPTCF"

ORIGIN

Alignment Scores:

Pred. No.: 2,84e-179 Length: 1152
Score: 2031.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-661-430-1 (1-383) x AY533306 (1-1152)

Qy 1 MetAlaPheArgHisLeuAlaValAlaArgLeuLeuSerLeuLeuValLeuCysAlaVal 20
Db 1 ATGGCTTTTCGTCATTTGGCAGTCGCCAGACTCACTGCTTGTCTGCTTGTGCGCGT 60
Qy 21 LeuLeuLeuValHisAlaMetIleTyrIleProSerLeuTyrGluAsnLeuThrIle 40
Db 61 CTTCATTATGATTCATGCATGATTTATAGATTCCTTACGAGAACCTTACTATAC 120
Qy 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGluValSerThrAla 60
Db 121 GGCTCCTTCACCTTATTCGCCAGCGCAGCAGAGAGAGAGAGTGTCTGGGAATACGAGCT 180
Qy 61 SerThrSerAspAspLeuLeuAspThrTyrPasnSerThrPheSerProIleSerGluVal 80
Db 181 TTCACCTGGAGATGATCTTATGATAGTGAATTCACGCTTTCCACCGATTTCTGAGATT 240
Qy 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
Db 241 AATCAGACTAGTTTATGAGAGACATTCCTCAATCTGTCCTCCGACCAACAGACTCTT 300
Qy 101 GlnPheCysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGlu 120
Db 301 CAATTCCTGTAATCAGACACCTCCCACTGTCGAGCCCATCGGTATTTCTCGATGAG 360
Qy 121 ProAspPheLeuSerThrLeuGluIleTyrIleTyrProAspThrHisAlaGlyGlyHisGlyMet 140
Db 361 CCCGACTTCAAACTCTCGAGAAATCTATCCGACACCCGCGGTGAGACATGAGATG 420
Qy 141 ProLeuAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGlu 160
Db 421 CTTAAGAGATTGTTGCCAAGCATCGTGTCTATATTGTCCTTATGAGATCGTGAA 480
Qy 161 AlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaIleGlnGlnLeuAsp 180
Db 481 GCACATTGAGAAATATGCTCCACATTTGCACTGCTGCGCAAAACAAACATTTGAGC 540
Qy 181 TyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyIleLeuMet 200
Db 541 TATGCAATTTTCATTTGTGAGCAAGTGGCGAATCAGACTTTATATGCGGGAACATATG 600
Qy 201 AsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAsp 220
Db 601 AACGTTGGATACGACATCAGCCCTCTACCCATGCGAGTGTTCATCTTTCATGAT 660
Qy 221 ValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHis 240
Db 661 GTGAGATTACTGCCGAGATGACCGTAACTGTACAGTGTCCATTCACACGCTCAT 720

Qy 241 MetSerValAlaIleAspLeuPheAsnTyrIleLeuProTyrSerAlaIlePheGlyIle 260
Db 721 ATGAGTGTAGCGATCGATTAATTCATTAATACTTCATATTCGGCGATCTTCGGCGGA 780
Qy 261 IleSerAlaLeuThrLysAspHisLeuLysIleLeuGlyPheSerAsnAspPheTyr 280
Db 781 ATCAGTGCATCAACAAAGATCACTGAAGAAATCAATGATGATTTTTCGATATTTTTCG 840
Qy 281 GlyTyrGlyIleGluAspAspLeuAlaThrArgTyrSerMetAlaGlyLeuLeuVal 300
Db 841 GGTGGGGGAGAGAGACGACGATTTGGCGACGAGAACATCGATGCGCTGACCTGAAAGTT 900
Qy 301 SerArgTyrProThrGlnIleAlaArgTyrIleMetIleYHisSerThrGluAlaThr 320
Db 901 TCAAGATATTCGACACCAATTCACAGATTAATAAGATTAGCATCTCGACGGAACGACG 960
Qy 321 AsnProValAsnLysCysArgTyrIleMetGlyGlnThrLysArgArgTyrThrArg 340
Db 961 AATCCAGTTAATAAATGCGCGCTACAAATATGAGCCCAACGACGCGATGACACGT 1020
Qy 341 AspGlyLeuSerAsnLeuLysTyrIleLeuValAsnLeuGluLeuLysProLeuTyrThr 360
Db 1021 GAGGCGCTTAGCAATCTGAATAGCTCGTAATCTGAAATTAAGCTCTTCACT 1080
Qy 361 ArgAlaValAlaAspLeuLeuGlyLysAspCysArgArgGluLeuArgAspPhePro 380
Db 1081 CGAGCGGTGCTGATTTGCTCGAATAAAGATCCGCCCGGAGCTCGAAGGACCTTCCA 1140
Qy 381 ThrCysPhe 383
Db 1141 ACGTGTCTT 1149

RESULT 3
AC025727/c 63747 bp DNA INV 27-MAY-2005
LOCUS AC025727
DEFINITION Caenorhabditis elegans cosmid Y73E7A, complete sequence.
ACCESSION AC025727
VERSION AC025727.3 GI:14530915
KEYWORDS HTG.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 63747)

REFERENCE
AUTHORS
CONSTRM
TITLE
JOURNAL
PUBMED
9851916
2 (bases 1 to 63747)
AUTHORS
Du,H. and Maupin,R.
TITLE
The sequence of C. elegans cosmid Y73E7A
JOURNAL
Unpublished (2001)
3 (bases 1 to 63747)
AUTHORS
Waterston,R.H.
TITLE
Direct Submission
JOURNAL
Submitted (12-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 63747)
REFERENCE
AUTHORS
Waterston,R.
TITLE
Direct Submission
JOURNAL
Submitted (13-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 63747)
REFERENCE
AUTHORS
Waterston,R.
TITLE
Direct Submission
JOURNAL
Submitted (23-JUN-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
6 (bases 1 to 63747)
REFERENCE
AUTHORS
Waterston,R.

TITLE Direct Submission
JOURNAL Submitted (03-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 7 (bases 1 to 63747)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 8 (bases 1 to 63747)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 9 (bases 1 to 63747)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 10 (bases 1 to 63747)
AUTHORS Wormbase Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 11 (bases 1 to 63747)
AUTHORS Wormbase Consortium
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 12 (bases 1 to 63747)
AUTHORS Wormbase Consortium
TITLE Direct Submission
JOURNAL Submitted (27-MAY-2005) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
COMMENT On Jun 23, 2001 this sequence version replaced gi:10800387.
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: submissions@watson.wustl.edu and jes@sanger.ac.uk
NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.
This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m3 subclone.
For a graphical representation of this clone sequence and its analysis see:
<http://www.wormbase.org/db/seq/sequence?name=Y73E7A;class=Sequence>
NEIGHBORING CLONE INFORMATION
The 5' clone is Y92H12A, 200 bp overlap; the 3' clone is Y71G12B, 4000 bp overlap.

NOTES:
Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yui Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFome cloning project (<http://worldb.dfci.harvard.edu/>), similarity to other proteins from Blastx analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's MABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).
Location/Qualifiers
1. 63747
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="Y73E7A"
5487..12773
/locus_tag="Y73E7A.3"
join(5487..5659,7957..8063,8778..8971,9035..9119,10020..10254,11408..11615,11668..11742,11836..11943,12669..12773)
/locus_tag="Y73E7A.3"
/standard_name="Y73E7A.3"
/note="contains similarity to Pfam domain PF06027 [bukaryotic protein of unknown function (DUF9141)]; coded for by the following C. elegans cDNAs: YK1142a05.3, OSTF30045C02"
/codon_start=1
/product="Hypothetical protein Y73E7A.3"
/protein_id="AAG23391.3"
/db_xref="GI:60543518"
/translation="MTNCGAESTSAPPQNGHLRPPGIOPHRNPPTPPSPDEPDDAINGSPCCDDPVRVTFKALIGQLSLGCGVSSQLLVNQNVAAPAAQAFNVELCPVYVTSIACKNDENGLVIVRKRMGRVLLIALLIDVENWYIVKAYQVNTLTSVOLLDCATPTVPLFLSVRYLASHITGVITCIIGACVYMAQDLGDKNSKYLGHILCAAAMVAICNVAEFLVKQSRREYVGMGLFECTVSGVTGVAPEQALSKIWMGDTIVSYFLAFASMPFIYSILVTVLQKTSALMNFSLTADFYSLLFGIFMFDOTFHLVYVSFIVLCIGSIVSKKETQVRADDEFRRVPCPLFVCCCGCFEEDGSTEGSTIVSPSPNERMOMGLNPRGSMPCPVHGSQQLSHL"
complement(14300..15495)
/locus_tag="Y73E7A.4"
complement(join(14300..14457,15222..15495))
/locus_tag="Y73E7A.4"
/standard_name="Y73E7A.4"
/note="contains similarity to Pfam domain PF05835 (Synaphin protein); coded for by the following C. elegans cDNAs: YK1215a10.5, YK1728h08.3, OSTF30035H07"
/codon_start=1
/product="Hypothetical protein Y73E7A.4"
/protein_id="AAG23387.1"
/db_xref="GI:10800392"
/db_xref="Wormbase:Y73E7A.4"
/translations="MAGFLMKQVNGQSLVYTGGLMKDGGKRTGDEPPEYIARLEQDERKEKRRKREKRRQGIIRVKAIVKKKEBVANVDFGRIGRPKTEETAAENNAEDDSLIGQLTEQVEKAKTMAVTAFEIVKGFPFGK"
complement(19944..20816)
/locus_tag="Y73E7A.5"
complement(join(19944..20130,20270..20541,20589..20816))
/locus_tag="Y73E7A.5"
/standard_name="Y73E7A.5"
/note="contains similarity to Pfam domain PF07735 (F-box associated)"
/codon_start=1
/product="Hypothetical protein Y73E7A.5"
/protein_id="AAG23386.1"

/db xref="GI:10800391"
 /db xref="MormBase:Y73E7A.5"
 /translation="MFOHIFIKFOIIGDIFRLSYOSALSHPLFRKCHFLKPLGEGSYI
 TTEDMKFLIKNPNLKYGLTRSLRRTKIDIKLVNIPUDPEKSGULDLDAKNMCE
 KMRFDLNTTKAREINKYIRYLDGNLPLNRRLQNMWHMHWAGDLGLPLSKNPK
 RRARLFCDLEFVDCSYGRDIERSDGLATLILNFAMVFLVMHDFPTDPTLPHFS
 NSVSVPTNN"
 21160..26092
 /locus_tag="Y73E7A.2"
 /locus(21160..21238,21285..21638,23801..24363,25697..26092)
 /locus_tag="Y73E7A.2"
 /standard_name="Y73E7A.2"
 /note="contains similarity to Rattus norvegicus AATP
 protein (Apoptosis antagonizing transcription factor).
 TR:Q9QYWO; coded for by the following C. elegans cDNAs:

Alignment Scores:

Pred. No.:	9,11e-107	Length:	63747
Score:	1275.00	Matches:	381
Percent Similarity:	22.50%	Conservative:	0
Best Local Similarity:	22.50%	Mismatches:	2
Query Match:	62.76%	Indels:	1312
DB:	2	Gaps:	5

US-10-661-430-1 (1-383) x AC025727 (1-63747)

QY	1	MetAlaPheArgHisLeuAlaValAlaArgLeuLysSerLeuLeuValLeuCysAlaVal	20
Db	45925	ATGGCTTTTCGATCTTGGCAGTCCGACAGCTCAAGTCGTGCTGACTTTGTGCGCT	45866
QY	21	LeuLeuLeuValHisAlaMetIleTyrLysIleProSerLeuTyrGluAsnLeuThrIle	40
Db	45865	CTTCATTAGTTCAGCAATGATTTATTAAGATTCATGCTTTACAGAGAACCTTACTATC	45806
QY	41	GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGluValThrAla	60
Db	45805	GGCTCCTCCAGCCCTTAATGCCACGTCGACGCAATGAGGACAGTGTCCGGAAATACGGCT	45746
QY	61	SerThrSerAspAspLeuLeuAspThrTPAAsnSerThrPheSerProIleSerGluVal	80
Db	45745	TCCACTTCGAGATGATCTACTGATACGTGAAATCCAGCTTTTCCAGATTTCTGAGTT	45686
QY	81	AsnGlnThrSerPheMetGlu	87
Db	45685	AATCAGACTAGTTTATGGA-GGGTAAGGTTCTTAACCTTATTTGGAATTTTGGAA	45627
QY	87	-----	87
Db	45626	AAAAAATTTTTCGAAATTTTGAAAAAATTTGAATTTTTCCTCCAAATCTCAAA	45567
QY	87	-----	87
Db	45566	ATTCCAGTGTGCTGACGACTGTCCTTTACGGTTGATCTACAAAAAATGCGGAAT	45507
QY	87	-----	87
Db	45506	TTTTTGGCCAAAAAATGTGACCTCAGCAGACAGTCTTAAACATGCAAAATCAGTGAAGA	45447
QY	87	-----	87
Db	45446	CTTCGGCCTCTTTCCTCCGCAATTTTGTAGATCGAGTAGATCAAGCGAAATGGGGG	45387
QY	87	-----	87
Db	45386	CTTTGAGACCGGTGTATGTAAAAAATTTCCAGAAATTTCAAAAAAATTTGCAA	45327
QY	87	-----	87
Db	45326	GATATAAAAAAGTTTTTTTCCAAATTTGAAAAAATCTGAATATTTCCAAATTTCTTGC	45267
QY	87	-----	87
Db	45266	AAAAATTTTCTTTTGAATTTTATTTGAAAAAATTCAAATTTCTGTTTATTT	45207

QY	87	-----	87
Db	45206	TTTGGATTTAAAAAATAATTAATGTCACAAATTTTCCACAAATATCCAAAAATTGG	45147
QY	87	-----	87
Db	45146	AATTTTCTCMAAAAAAAGAAATTAATTCMAATTGTCCAGTGAATACACAAT	45087
QY	87	-----	87
Db	45086	CGTCTTTTCTTTTCAATTTGGCCAGCCGCAAAAAATGCTAATCAGAAATTTTGGG	45027
QY	87	-----	87
Db	45026	GAAATTTGAAAAATTTGAATTTTAAAAAATTAATAATTTTCTGTAAAAAATTA	44967
QY	87	-----	87
Db	44966	AATTTCTTTTAAAAAAGCAGCAAAACAGACAAATTTTACAGGATCTGACTTCCT	44907
QY	87	-----	87
Db	44906	CACATGTTGAATGAGAGATTGGCCGAACTAGGCCAGGCCATATCTGGGTGATTAC	44847
QY	87	-----	87
Db	44846	GGCGGATTTACGGCGCACTACTTTCCAGTCCGACAGCGCATTTGCAATGAGGCGGA	44787
QY	87	-----	87
Db	44786	AAATTCATGAGAGAGCCAGAACCCCATGATTTCTAATTAATTAATAATTTGAAAA	44727
QY	87	-----	87
Db	44726	AAAACAATTGAGATTGAGAAAAATTTTGAATAATTCAGAAAAAATTTGTGTAATTT	44667
QY	87	-----	87
Db	44666	TGAATTTTATTTAAGGAAACACAAATTTCTGAGAAATGCGTACTGACACATATTTGA	44607
QY	87	-----	87
Db	44606	CGCGAAATATCTGTAGCAAAAACTGCAATCATTTGACTAGTGGCTGTGTGCAC	44547
QY	87	-----	87
Db	44546	TTACGGCCTCGTTGATACAGTATGATTAAGAAATTAATCTGATGTTTCGCTACGAG	44487
QY	87	-----	87
Db	44486	ATATTTGGCGGTCAATATGTTGTGTAATACGCAATTCGAAAAATTTTCTGTCCATTA	44427
QY	87	-----	87
Db	44426	TAAAGTAAAAAATTAATTAATTTATCTTTAGTAAAGCAATTTTATAGCGTAAGATT	44367
QY	88	-----	87
Db	44366	GAGCATTAATACCAATTTTCAAAAAAATCCAAATTTCCAGAAATTCGTAATTCCTG	44307
QY	94	PheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyPro	113
Db	44306	TTCCCGACAAACAGACTCTTCAATCTGTATACGACACCTCCCACTGTGTGGACCC	44247
QY	114	IleArgValPheLeuAspGluProAspPheIleThrLeuGluLysIleTyrProAspThr	133
Db	44246	ATCCGTGATTTCTGATGAGCCCGCACTTCAAACTCTCGAAGAAATCTATCCGACACG	44187
QY	134	HisAlaGlyValHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle	153
Db	44186	CAGCGCGGTGACATGGAATGCTTAAGGATTTGTGTGAAGGCATCGTGTGCTATTAAT	44127

QY 154 ValProTyrArgAspArgIuAlaHisIleu----- 163
Db 44126 GTCCTTAATAGAAATGCTGAGACATTTAGGTAGAAATTCGAATGCTGGGCTCT 44067
QY 163 ----- 163
Db 44066 GGCACGATTACGGAATTTGAGAAATTTGAAAAAAATTTGAAAAAAATTTT 44007
QY 163 ----- 163
Db 44006 TAAATAATTTTGAAATTTTAAAAATCAAAATTTGAACTTGAGATCAAAAAACAT 43947
QY 163 ----- 163
Db 43946 GTTAGGTATCATTTGGGTCTGGCTGAATTAACGAATTTTCAGAAATTTGAAACTTTT 43887
QY 163 ----- 163
Db 43886 TTGGAAACGAAAAAATTTGTGAAATTTGTAAATTTTTCGATTTTAAAAA 43827
QY 163 ----- 163
Db 43826 TTTCAAAATTTCAAAATCCCAAACTGTCAATTTTGTCCAAAAAAGAGTAATCCA 43767
QY 163 ----- 163
Db 43766 CAATTCATCCGCGCAATCGACCTTCCCTATAAGTGAATTTTGTCCGGGATTT 43707
QY 163 ----- 163
Db 43706 CCAACTACGCACTTTTCCACTGAATTTTGGTTTTAGACATTTTGTAGATAAAT 43647
QY 163 ----- 163
Db 43646 TGAAGCTCAGAACGAGATTAAGAAATTTAGGAATTTAGTCTCGGACGATTT 43587
QY 163 ----- 163
Db 43586 ATAAAGTTTGAAAAAATTTGAAAAAGAAATTTTGTGAAAAATTTGAAAAAAGTTAA 43527
QY 163 ----- 163
Db 43526 AAAAAATTTTGTGATTAATTTCCATATTTATCATAGTAATATATACCTTTTC 43467
QY 163 ----- 163
Db 43466 AAAAAATTCCTCACTCACTGCTGCTCAAGTGTCTCATTTGATCTACGTTGATCTA 43407
QY 163 ----- 163
Db 43406 CAAAAATGGGGGAGAGGGGCAAGAGTTCTCAAGTATTCGATGTTAGAAAGCTG 43347
QY 163 ----- 163
Db 43346 CTGACGTCAATTTTGTGGCAAAAAATTCGCGCATTTTGTGATCAAAATCCTGAT 43287
QY 163 ----- 163
Db 43286 GAGACAGCTGGACCACTGTGCTCAATGGAATTCGAAACCCGTAATTTCAACTACCCC 43227
QY 164 ----- 174
Db 43226 GATTTTCGATTCAGCTATATTTTTCAGAAATATGCTCCGCAATTTGCACTGCTCTC 43167
QY 175 AlAlaIysGlnIleuAspTyrAlaIlePheIleValGlnIleValAlaAsnGlnThrPhe 194
Db 43166 GCCAAACAAACAATTTGACTATGATTTTCATGTGAGCAAGGGGAAATCAGACGTTT 43107
QY 195 AsnArgGlyIys----- 198
Db 43106 AATCGCGGAAAGTATGTTATTTATTTATTTATTTTCTCCAAATTTCTAAACAATTGGA 43047
QY 198 ----- 198

Db 43046 GAAAAAATGATATATATCTTTGGGGTCCCGGTGACGTCACGCGTGGCGGTGAAAAAA 42987
QY 198 ----- 198
Db 42986 ATCAAAAATATAGGTAGGGGAAACCCACGTGTGTACAGCTGTCCCAATACAGTCTGATCT 42927
QY 198 ----- 198
Db 42926 ACAAAAAATGCGGAATTTTTCCAAATTTGTAGCTCATCAGCTCTTAATTATGCANA 42867
QY 198 ----- 198
Db 42866 ATCAGTTGAAACTGTGCGTCACTATTTTTCGAAGATCAAAAGCAATGGACTTTT 42807
QY 198 ----- 198
Db 42806 TGACTCCAGTGGACCCCTGTGACGTCAAGGTGGCCGTGCTTAAAAATTCAAAAAT 42747
QY 198 ----- 198
Db 42746 AGAGTTGGGGACCCATCTGACGTCAATGTGACATCATAGAGGCGCGCGGCTTGA 42687
QY 199 ----- 209
Db 42686 ATCTCCAAATATAAAAATTTCTTCAGCTAATGAAGTTGATACGACGTAGCATCACGC 42627
QY 210 LeuTyrProTyrGlnCysPheIlePheHisAspValAspLeuProGluAspAspArg 229
Db 42626 CTCTACCAATGGCAGGTCTTCAATGATGATGATTTCTGCCCCAAGATGACCGT 42567
QY 230 AsnLeuTyrThrCysProIleGlnProArgHisIleMetSerValAlaIleAspIysPheAsn 249
Db 42566 AACCTGTACCGGTGTCCATTTCAACCACTCATATAGTGTAGCGATTCATTAATTCAT 42507
QY 250 TyrIysLeu----- 252
Db 42506 TATTAAGTA-GGTTTTTCTGATAAAAAATTTCAAAAAATCTGAAAAAATCGATTCC 42448
QY 252 ----- 252
Db 42447 AACTTTCAAGTCTGGCGCTTCTCATTTGAAATTTTTCGGCTCAATTGACAAATCCCGC 42388
QY 252 ----- 252
Db 42387 CAGGACACGCGGTAGTGTGTACTCCACAGGACAAATACATTTAGTTTGCAAC 42328
QY 252 ----- 252
Db 42327 TAAATCGAGCGCGACGACGACGACGCGCGTAAATCTATGAGCCAAATAGCCT 42268
QY 252 ----- 252
Db 42267 AGTTGGCGAAAATCTTCCATTTTAATTATGAGGAAAGCAATGATTGAATG 42208
QY 252 ----- 252
Db 42207 AAGATTTTGTGAAGTACGCAATTTGGCTGGGCAATATCTGATGATTTACGGCG 42148
QY 252 ----- 252
Db 42147 CGTTGCGTGGCGGTGCGGTTGATTTTACTTTGTAATCTAATGATTTGTGCGGTG 42088
QY 252 ----- 252
Db 42087 GAGTACACGACACTTTCACACGCGTTGTCGCGCGGCGAATGAGCGGAAAAA 42028
QY 252 ----- 252
Db 42027 TTCAATGAGAGGCGCGAACCCTGTTTAATATCATTTTTCGAAAAAATACAAAA 41968
QY 253 ----- 262
ProTyrSerAlaIlePheGlyIleSer 262

```

Db 41967 AATTGAGATTAACTTTTTCAGACTTCATATTCGGCGATCTTCGGCAATCAGT 41908
Qy 263 AAlaLeuThrLysAspHisIleuLeuYelYsIIleAsnGlyPheSerAspPheThrGlyTrp 282
Db 41907 GCACTAACAAAGACCTGCTGAAGAAATCAATGATTTTTCAGATGATTTTGGGGCTGG 41848
Qy 283 GlyGlyGluAspAspAspleuAlaThrArgThrSerMetAlaGlyLeuLeuValSerArg 302
Db 41847 GCGGAGAGAGACGACGATTTGGCGAGACATCATGATGCTGCGACTGAATTTTCAGA 41788
Qy 303 TyrProThrGlnIleAlaArgTyrLysMetIleYsHISerThrGluAlaThrAsnPro 322
Db 41787 TATCCGACACAAATTCACGATATATAATGATTAAGCACTCGACGGAAGCGACGATCCA 41728
Qy 323 ValAsnLys----- 325
Db 41727 GTTAATTAAGTACGAGAAAGAAATATATCGATTTTTCGCGCAAAATTTGTCGAAATCCGA 41668
Qy 325 ----- 325
Db 41667 TGAATACTACGATTTTGTGAAAAAATCGAAAAAACAACAAAAAATCGATATAT 41608
Qy 325 ----- 325
Db 41607 TTTTGTAGAAATTCGATTTTAAAGAAATCGATATATTTATGTTTTCGAGAA 41548
Qy 325 ----- 325
Db 41547 AAATCGATTAATTTATGATTTTCGAAAAAATCGATTAATTTATGATTTTAAAGAAA 41488
Qy 325 ----- 325
Db 41487 AATCGATTAATTTATGATTTTAAAGAAAAATCGATATTTATGATTTTGAATTA 41428
Qy 325 ----- 325
Db 41427 AAAAGCATTAATTTATAGTTTATTTTAAATTAATCGATTAATTTATGATTTT 41368
Qy 325 ----- 325
Db 41367 TTCGTAAAAAGTCGATGACTTTATGATTTTGAAGAAAAAATCAATATTTTCTTT 41308
Qy 325 ----- 325
Db 41307 GATTATGAGAAAAATCGATATTTTATTTATTTTTCGTAAGAAAGTCGATGACTTT 41248
Qy 325 ----- 325
Db 41247 TATGATTTTTCGAAAAAATCAATTAATTTTCTTATTTTTCGAGAAAAATCGAT 41188
Qy 325 ----- 325
Db 41187 AATTTTATGATTTTTCGAAAAAATCAATTAATTTTTCGATGAGAAATCGATTTTCTT 41128
Qy 325 ----- 325
Db 41127 AAAATCGATTAATTTATAGATCTTCGAGAAAAAAGTAGAGATTTTATGATTTTAAAG 41068
Qy 326 ----- 330
Db 41067 CTAAATAATCTAAATTTTTCACCTAAATAATTTTCGTTCCAGATGCGGCTACAAATA 41008
Qy 331 MetGlyGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLeuYelYsLeu 350
Db 41007 ATGGGACCAAAACGACGCGGATGACACGTCGACGCTTAAGCAATCTGAAGATTAAGCTC 40948
Qy 351 ValAsnLeuGlnLeuLeuYsProLeuYrThrArgAlaValValAspLeuLeuGlnYsAsp 370
Db 40947 GTAATCTCGAATTAAGCTCTCTACACTCGAGCGCTGCTGATTTGCTCGAAAAAGAC 40888
Qy 371 CysArgArgGlnLeuArgArgAspPheProThrCysPhe 383
Db 40887 TGCCCGCGGAGCTCGAAGGACCTTTCACACGTGTTT 40849

```

```

RESULT 4
LOCUS AY601103 1269 bp mRNA linear INV 02-AUG-2004
DEFINITION Trichoplusia ni beta 1,4-N-acetylglactosaminyltransferase mRNA,
complete cds.
ACCESSION AY601103
VERSION AY601103.1 GI:47156062
KEYWORDS
SOURCE Trichoplusia ni (cabbage looper)
ORGANISM Trichoplusia ni
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Plusiinae; Trichoplusia.
REFERENCE 1 (bases 1 to 1269)
AUTHORS Vadaie,N. and Jarvie,D.L.
TITLE Molecular Cloning and Functional Characterization of a Lepidopteran
Insect (beta)4-N-Acetylglactosaminyltransferase with Broad
Substrate Specificity, a Functional Role in Glycoprotein
Biosynthesis, and a Potential Functional Role in Glycolipid
Biosynthesis
JOURNAL J. Biol. Chem. 279 (32), 33501-33518 (2004)
PUBMED 15173167
REFERENCES 2 (bases 1 to 1269)
AUTHORS Vadaie,N. and Jarvie,D.L.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2004) Molecular Biology, University of Wyoming,
Laramie, WY 82071, USA
FEATURES
source location/Qualifiers
1..1269
/organism="Trichoplusia ni"
/mol_type="mRNA"
/db_xref="taxon:7111"
1..1266
/note="b4GalNCT; glycosyltransferase"
/codon_start=1
/product="beta 1,4-N-acetylglactosaminyltransferase"
/protein_id="AA11926.1"
/db_xref="GI:47156063"
/translation="MGRATRAARLLLLVLAALAVELRSGSIDASPLRYLYTPY
NATQPTLNVRLANWPKI PSNVIDSEBSYISKISLSNHTTRASVHPSSITET
ASKDKMNTIDOGAFAMISPTPLITKIMDSISKYVTTEDVKAALAVVLP,CDMP
PDLGPTLNTKTELELEWEKKEPVEWGRSPNCTARHVAIIVPRDQOHAIIF
LNHNPFLMKQOIEYGLPIVEQGNKDPRAKLNWGVESOKIYVAEMOCQFVIDID
LLPLDTRNLVSCPRQPRMSASIDKIHKKIPEYDIFGVSAITLQFPNVNGFSKWK
GWGSDDDMSYRLKINHIARIMKSIARIMLHKSTPNPKRYQLSOTSKITPKO
GLSTLEYELVGVVHYHLTVILVINDERS"
ORIGIN
Alignment Scores:
Pred. No.: 2,28e-62 Length: 1269
Score: 775.50 Matches: 157
Percent Similarity: 60.94% Conservative: 63
Best Local Similarity: 43.49% Mismatches: 116
Query Match: 38.18% Indels: 25
DB: 2 Gaps: 6
US-10-661-430-1 (1-383) x AY601103 (1-1269)
Qy 30 LysIleProSerLeuYr-----GluAsnLeuThrIle 40
Db 190 AAGATACCTCAATTAATACATGAGAACTCCGAGAAGATCTCAATTAATAAATCAATTGCTT 249
Qy 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla 60
Db 250 AGCAATCACACGACAAAGACGAGCTGTACACCCGCGCAGACAGTATACGAGACTCG 309
Qy 61 SerThrSerAspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluVal 80
Db 310 TCGAAACTCGATTAACAATATGACATTCAGAGTGTGCTTCGCTATATATGAGCCCAACA 369
Qy 81 AsnGln-----ThrSerPheMetGluAspIleArgProIleLeuPheProAspAsn--- 97

```


gene

/note="Vector: PCMV-SPOrT6.1"

1..2265

/gene="bagalt3-prov"

/note="synonym: MG084507"

/db_xref="GeneID:446718"

178..1188

/gene="bagalt3-prov"

/codon_start=1

/product="bagalt3-prov protein"

/protein_id="AAH77601.1"

/db_xref="GI:50417573"

/db_xref="GeneID:446718"

/translation="MALVIREGGRKVAYPEGIFVGSGASMGILPHNLSQADGVY
NLSLISKPSIREELPYCEPTSPFPGPIRVKLPEULLBEVRKANPVTGGRYRKE
DESEHTKTAVALI PHNRREQHLKYLLHPFLQRQLNGIYIIHQAGETTFARAKI
NVGFEXAMDEDMDICLFPHDVID.IPEDDDNIYCDEFPKIASIAMPFGLYKSKS
GSVASLPEDQMKNGFPNNYWGMSGEDDIIGRAVALSGMLISRPVSQYRGYMIKH
HDKNQEONPKRFNMVLTKRRITMKQDMNLQYLISKELPLTYNTIVDIGDKLR
HT"

CDS

ORIGIN

Alignment Scores:

Pred. No.:	7..56e-59	Length:	2265
Score:	741.00	Matches:	145
Percent Similarity:	59.95%	Conservative:	78
Best Local Similarity:	38.98%	Mismatches:	115
Query Match:	36.48%	Indels:	34
DB:	5	Gaps:	6

US-10-661-430-1 (1-383) x BC077601 (1-2265)

Oy 4 ArghisLeuValaValaAargLeuLySerLeuLeuValaLeuCyValaValaLeuLeu 23

Db 115 AAAAAATGTATCTCTTAAGGGCGCTACCTGTTACTTCCTGCAGCAGTAGTG 174

Oy 24 ValHisAlaMetLefyr-----LytleleProserLeutyrgLiAsn 37

Db 175 GTCATGCCCTCTGTTATAGAGAAGGCTACAGAAAAGGCGCTATTGTTGGGATA 234

Oy 38 LeutThrilegLySerSerThrlenuleialaasPvalaAspIlaMetGiualaValleugly 57

Db 235 TTTTCAAAAGGGGGCGTCTCT-----TCCATGGAGCATCAACCCC 276

Oy 58 AenThrAlaSerThrSerSeraspAspLeuAspPthrTTPasnSerThrPheSerProile 77

Db 277 CACAATTATCCCAAGCTGGGAGACGTT-----TATGCCAACCTC 315

Oy 78 SerGiUValaAsnGlnThrSerPheMetGiUasPilleArgProileleuePheProaspasn 97

Db 316 ACCTTAATATCCAAAGCGCTTCGCGGGA----- 345

Oy 98 GlnThrleuGlnPheCySaSnGlnThrProProHiSleuValGlyProileArsvallPhe 117

Db 346 GAGGAGCTCCCATCTCTGCTGAGACCCTCCCTTTTACGGGGGCCAAATPAAGATCAAG 405

Oy 118 LeuAspGiUProAspPheLyStrhleuGluysile-----TyProAspThrhis 134

Db 406 TTG-----CCGGAGAACTTGTAACGGAGGAAGTGAACGCAAAATCCTTACGTGACC 455

Oy 135 AlaGlyGlyHisGlyMetProLyAspAsyValAlalaArgHisarGValAlaalleleval 154

Db 460 AAAGGGGGCGGTACAAAGCGCGCCGAGCTGCAACCATMAAATCGCCGATCATC 519

Oy 155 ProTyrrArgAspArgGluAlahlsleuArgIlametleuhIsanleuihsaserleuLeu 174

Db 520 CCTCACGGCAACGGGAGACGACCTTAATATATGTGCTCTACACTTGCCACCCGTTTTCTG 575

Oy 175 AlalyegInGlnleuAspTyrralalepheileValGluGlnValalaasenglnThrPhe 194

Db 580 CAGCGGCACCACTCAACTACGGGATATATCATCATCATCAGGAGGTAAATTCACATTC 639

Oy 195 AsnaArgGlyLyuleuMetAsnValIGlyTyrrAspValaIaseArgleutyrrProtPrGln 214

Db	640	AACGAGACGAAGTTACTGAAAGTCGCGCTTCAAGAGAGCCATTGAAAGACGAGACTGGAC	699
Oy	215	CysPhe11ePheh1saapValaapLeuLeuPProGluaspAspArgaenLeuYrThrCys	234
Db	700	TGCCTGTTCTTCCACGATGTCGACCTTCATCCCGAGAGACATGTAACATTTATATATGT	759
Oy	235	Pro11e1InProaThgH1smetSerValA11a1aaspAspPheaaenYrLyb1eLeuProTy	254
Db	760	GACAGATTTTCCAAAGCAGCCTCCATCGATCCATGGACAGATTGGATACAAATTGCTTAC	819
Oy	255	SerA11a1ePheG1yG1y11eSerA1aLeuThrLybAspH1sLeuYsLy11eAaG1y	274
Db	820	AAATCTCACTTCGAGAGAGTGTCTGCCCTGAGCCACAGCAATACATGAAGATGAATGA	879
Oy	275	PheSerAaaspPheThrPglYrTPG1yG1yG1uaspAspAspLeuA1aThrArgThrSer	294
Db	880	TTTCCCAACAACATCTAGCTGGGGCTGGGGGGGTGAAGATGACACATAGGATCAGAGTGGCA	939
Oy	295	MetA1a1yLeuLeuYsValSerA1gYrYrProThG1n1a1aArgYrLybSmel11eLy	314
Db	940	CTCAGTGGCATGTCTTAATCTCGCCCTTCAGTCCAGACGAGAGATATAAAATGATCAAG	999
Oy	315	H1serThrG1uA1aThraaPProValAa1aSmLyCysArgThrYrLyb11eMetG1y1nThr	334
Db	1000	CACGGCCATGACAAAGGCAATGACCAAAATCCCAAAAGTTCAACATGTGACTAAGACT	1059
Oy	335	LybArgArgYrThrArgaapG1yLeuSerAaenLeuYsYrLybLeuValAaenLeuG1u	354
Db	1060	CGAGAGCATTGGAAGCAAGATGGAAATGAACGCACTGACATGATACATGATCTCCAAGAG	1119
Oy	355	LeuYsProLeuYrYrThrArgA1aValA1aSpLeu	366
Db	1120	CTGCAGCCACTATACACCAATATTAACCGTGACATT	1155
RESULT 6			
LOCUS	CR942597	2098 bp	mRNA linear VRT 18-Apr-2005
DEFINITION	Xenopus tropicalis finished cDNA, clone Tga5115g12.		
ACCESSION	CR942597		
VERSION	CR942597.1	GI:62732412	
KEYWORDS			
SOURCE	Xenopus tropicalis (Silurana tropicalis)		
ORGANISM	Xenopus tropicalis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;		
	Xenopodinae; Xenopus; Silurana.		
	1 (bases 1 to 2098).		
AUTHORS	Amaya,E., Aehurst,J.L., Bonfield,J.K., Croning,M.D.R., Davies,R.M.,		
	Francis,M.D., Garrett,N., Gilchrist,M.J., Grafham,D.V.,		
	McLaren,S.R., Papalopulu,N., Rogers,J., Smith,J.C., Taylor,R.G.,		
	Voigt,J. and Zorn,A.M.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-APR-2005) Sanger Institute, Hinxton, Cambridgeshire,		
COMMENT	CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk		
FEATURES	Sanger Xenopus tropicalis EST/cDNA project.		
	This sequence is from a Xenopus Gene Collection (XGC) library, from		
	a library constructed by Aaron M. Zorn. cDNA was prepared from RNA		
	extracted from gastrula embryos. EcoRI-NciI cut cDNA was then		
ORIGIN	ligated into pCS107 with EcoRI at the 5' end and NciI at the 3'		
	end.		
	Vector: pCS107; Site_1: EcoRI; Site_2: NciI		
	Host: Escherichia coli XL1-blue.		
FEATURES	Location/Qualifiers		
	1..2098		
	/organism="Xenopus tropicalis"		
	/mol_type="mRNA"		
ORIGIN	/db_xref="taxon:8364"		
	/clone="Tga5115g12"		
	/clone_1db="XGC-gastrula"		
	/dev_stage="gastrula (stage 10.5-13 mixed)"		
Alignment Scores:			

```

Pred. No.: 8,53e-59 Length: 2098
Score: 740.00 Matches: 145
Percent Similarity: 60.22% Conservative: 76
Best Local Similarity: 39.51% Mismatches: 118
Query Match: 36.44% Indels: 28
DB: 5 Gaps: 5

US-10-661-430-1 (1-383) x CR942597 (1-2098)

QY 6 LeuAlaValAlaArgLeuYsSerLeuLeuValLeuCyAlaValLeuLeuValHis 25
   ::::::::::::::::::::
DB 13 ATGCTCATCGCTAGCGGTCTACCTTCTCTCTGCGCAGCTGGTGTCATG 72
   ::::::::::::::::::::
QY 26 AlAmetIleTyrlsIleProSerLeuTyrluAsnLeuThrIleGlySerSerThrLeu 45
   ::::::::::::::::::::
DB 73 GCTCTGCTGTATAG-----GAGGCTACAGAAAAG 105
   ::::::::::::::::::::
QY 46 IleAlaSprValAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAsp 65
   ::::::::::::::::::::
DB 106 GTGGCT-----TATTTCTTGGTATATTTTACAAAGGGGACCTGCTCCATGGA 156
   ::::::::::::::::::::
QY 66 LeuLeuAspThrTrpAn-----SerThrSerProIleSerGluVal 80
   ::::::::::::::::::::
DB 157 CTGATGAAACCCCAATGATATCCAGCCGGGACGTTTACGCCAAGCTCAGCTTATAT 216
   ::::::::::::::::::::
QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAsnGlnThrLeu 100
   ::::::::::::::::::::
DB 217 TCCAGCTTCTTACGGGA-----GAGGACTTG 246
   ::::::::::::::::::::
QY 101 GlnPheCysAnglnThrProProHisLeuValGlyProIleArgValPheLeuAspGlu 120
   ::::::::::::::::::::
DB 247 CCATTTCTGGCGGAGACCTCATATATAGGGGGCCCAATAGAGTCAAGTTCGCCGAG 306
   ::::::::::::::::::::
QY 121 Pro---AspPheYsThrLeuGluValIleTyrlProAspThrHisAlaGlyHisGly 139
   ::::::::::::::::::::
DB 307 AACTTGTACTGGAGGAGGTGAGAAAAGAACCTTACGCTGTCCAAAGAGGGCGCTAC 366
   ::::::::::::::::::::
QY 140 MetProYsAspCysValAlaArgHisArgValAlaIleIleValProTyrlArgAspArg 159
   ::::::::::::::::::::
DB 367 AAGCCACCCGACCTCGAGTGCACATCAAAAGCGCGGTGATCATCCCTCACCGCGAGG 426
   ::::::::::::::::::::
QY 160 GlnAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnGlnLeu 179
   ::::::::::::::::::::
DB 427 GAGGAGACCTTAAATACCTGCTGTATATCTGCACCCGTTTCTGCGCGCGCAACTG 486
   ::::::::::::::::::::
QY 180 AspTyrlTrpAlaIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLeu 199
   ::::::::::::::::::::
DB 487 AACTACGGGATATCATCATCATCAGCGGGTAAATTTTACTTTCACCGCGCTAAGTGG 546
   ::::::::::::::::::::
QY 200 MetAsnValGlyTyrlAspValAlaSerArgLeuTyrlProTyrlGlnCysPheIlePheHis 219
   ::::::::::::::::::::
DB 547 CTGAACCTCGGCTTCAAGAGAGCCATGAAGAGAGATTGGAGCTGCTGTTTACAC 606
   ::::::::::::::::::::
QY 220 AspValAspLeuLeuProGluAspAspArgAsnLeuTyrlCysProIleGlnProArg 239
   ::::::::::::::::::::
DB 607 GATGTCAACCTCATTTCCGAAAGAGATCGCAATATCTTCACTCGCAAAATTTCCCAAG 666
   ::::::::::::::::::::
QY 240 HisMetSerValAlaIleAspTyrlPheAsnTyrlCysLeuProTyrlSerAlaIlePheGly 259
   ::::::::::::::::::::
DB 667 CACCCCTCGATCGGCATGAGCAAAATTTGGATACAAATGCTTACTTCTGCGG 726
   ::::::::::::::::::::
QY 260 GlyIleSerAlaLeuThrIleAspHisIleLeuYsIleAsnGlyPheSerAsnAspPhe 279
   ::::::::::::::::::::
DB 727 GAGGTGTACGCCCTTTCCCGAACAATACATGAAAGTGAAGCGCTTCCCAACAATAC 786
   ::::::::::::::::::::
QY 280 TrpGlyTyrlGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuYs 299
   ::::::::::::::::::::
DB 787 TGGGGATGGGGCGCGGAAGATGATGATATAGCATTCAGGGTCCGACTTACGCGCATGATC 846
   ::::::::::::::::::::
QY 300 ValSerArgTyrlProThrGlnIleAlaArgTyrlCysMetIleYsHisSerThrGluAla 319
   ::::::::::::::::::::
DB 847 ATCTCTGCGCTTCTATTCAGCACCGGAGGTATAAATGATCAACGAGCGCATGACAAAG 906

```

```

QY 320 ThrAsnProValAsnLeuCysArgTyrlsIleMetGlyGlnThrIysArgTrpThr 339
   ::::::::::::::::::::
DB 907 GGCATATGACCAAAATCCAAAAGTTTATACATGACGAAAGACCGCGGAGCGCG 966
   ::::::::::::::::::::
QY 340 ArgAspGlyLeuSerAsnLeuYsTyrlCysValAsnLeuGlnLeuYsProLeuTyrl 359
   ::::::::::::::::::::
DB 967 CAGGAGGGATGAATACCTGTCGAGTACCTATGCTTCCAAAGAGTTCGACGCGCTTAC 1026
   ::::::::::::::::::::
QY 360 ThrArgAlaValAspLeu 366
   ::::::::::::::::::::
DB 1027 ACCAATATCATCTGTGAACATC 1047
   ::::::::::::::::::::

RESULT 7
BC068719
LOCUS
DEFINITION
Xenopus laevis hypothetical protein MGC81163, mRNA (cDNA clone
MGC:81163 IMAGE:618052), complete cds.
ACCESSION
BC068719
VERSION
BC068719.1 GI:46250085
SOURCE
MGC.
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE
1 (bases 1 to 2681)
Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W.
and Richardson,P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
Dev. Dyn. 225 (4), 384-391 (2002)
JOURNAL
PUBMED
12454917
AUTHORS
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,K.H., Shennan,C.M., Schler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,D., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Umed,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mulhany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettelman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Sanchez,A., Whitting,M., Madan,A., Touchman,J.W., Green,E.D.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Scherer,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
PUBMED
12477932
AUTHORS
Klein,S. and Strausberg,R.
DIRECT SUBMISSION
Submitted (02-APR-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Igor David
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@gscc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth

```

REMARK
COMMENT

Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilia Prabu, Parraneh Seedat, JR Santos, Angelique Schermer, Ursula Skalska, Duane Smallus, Jeff Stott, Miranda Teal, George Yang, Jacqueline Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM: at: <http://image.llnl.gov>
Series: IRAX Plate: 156 Row: 1 Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source

1. 2681

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="MGC:81163 IMAGE:6318052"

/issue_type="Ovary, Xenopus"

/clone_lib="NICHD_XGC_Ov1"

/lab_host="PH10B"

/note="Vector: pCMV-SPORT6"

1. 2681

/gene="MGC81163"

/db_xref="GeneID:414540"

77. 1573

/gene="MGC81163"

/codon_start=1

/product="MGC81163 protein"

/protein_id="AAH68719.1"

/db_xref="GI:46250086"

/db_xref="GeneID:414540"

/translation="MMKRTLEKPCPLAMIGFQAPAMVYFSGRLSLTFKSHIEP
SDYKSTHDVSNLSVNSWVLPSTAEILIPCEPSYLVGPIVTSQPSLKK
VODKPLVKGGEYKPCNCAHRTAVIIPIHNRBTMLHLVLPLOKQOLYRI
YIHQAGNSTFNRAKLNVGVEALRDEMDCLFLADVDLIPENFNLVCCPMSPKH
ASVANKFSEYLYPYMPYFGVSGALTPDOYKKNAGFENYMGWGGDDIATVRLAGM
KITRPSVAVGHKMKHKGHDHNEPNHFDLITROMMKTKDGMSNSLYKLAELI
HLTYNTVDIGIDPRAKRLRNKRGPPRPIRYNKTSTSDSQIRNADQCLTVSE
INGVPSAEKKGPVAPBSTTTEKSLPIASAPRERVADKQIRVIRSTEDNNNAQV
KEVQADVARRGARRSSKLERQVDSFRAPKPDVKSILPQDLSRAPV"

ORIGIN

Alignment Scores:

Pred. No.:	2.44e-58	Length:	2681
Score:	736.50	Matches:	149
Percent Similarity:	60.78%	Conservative:	54
Best Local Similarity:	44.61%	Mismatches:	114
Query Match:	36.26%	Indels:	17
DB:	5	Gaps:	6

US-10-661-430-1 (1-383) x BC068719 (1-2681)

QY 55 ValLeuGLyAsnThrAlaSerThrSerAspLeuLeuAspThrTrpAsnSerThrPhe 74
DB 185 ATCTTCGGACACTCCATTGAGCCAGTTTGACTATTCGACAC---CATGATGCTAC 241
QY 75 SerProILeSerGluValaEngln-----ThSerPheMetGluAspIleArg 90
DB 242 ACCAACCTCTCCCTGCAACTCCATGCTGCTCCCTACTAGACACAGCAAGAGACATTC--- 298
QY 91 ProILeUhpPheProAspAsnGlnInThrLeuGlnPheCysAsnGlnInThrProProHisLeu 110
DB 299 -----CTGCTTCTCTAGATGCGCCAGAGAGGTCACCTTATTTG 334
QY 111 ValGlyProILeArgVal---PheLeuAspGluProAspPheLeuThrLeuGluIlySile 129
DB 335 GTGGGCGCCATTACCTTCACCTTCACCAACCCCTCCCTTAAAGAGTTCAAGTAA 394
QY 130 TyrProAspThrHisAlaGlyGlyIsglyMetProIlyAspCysValAlaArgHisArg 149
DB 395 AACCCCTGTGTATTAAGGTGGGAGTACAAGCCCGAAGTGTGAGGCGCGGACCCG 454

QY 150 ValAlaIleIleValProTyrArgAspArgIuaIaHisIleuArgIleMetLeuHisAsn 169
DB 455 ACGGCGGTCATCATCTCCACCGGAAACCGAGACTCACTGAGCACCCTGCTTATTAC 514
QY 170 LeuHisSerLeuLeuAlaLysGlnInLeuAspTyrAlaIlePheIleValGluIVal 189
DB 515 CTGACACCCCTTCCTGACCGGACGAGCTCCACTACAGATCTCATCATCATCAGGCG 574
QY 190 AlaAsnGlnInThrPheAsnArgIlyLysLeuMetAsnValGlyTyrAspValAlaSerArg 209
DB 575 GGTAACCTCACTTAATTAACCGGCGCAAGCTGTAATGTGGGCTGAAGAACTCTGGCG 634
QY 210 LeuTyrProTyrGlnCysPheIlePheHisAspValAspLeuLeuProGluLysAspArg 229
DB 635 GACGAGAGACTGGAGACTGTCTGTTTCTCATGATGATGATGATGATCTGAGATATCTTC 694
QY 230 AsnLeuTyrTrhCys---ProIleGlnProArgHisMetSerValAlaIleAspLysPhe 248
DB 695 AACTGTATGTGTGTGACCCCTGAGGCCCAACAGCGCTGTGATGATGATGATGATGATTC 754
QY 249 AsnTyrLysLeuProTyrSerAlaIlePheGlyIlyIleSerAlaLeuThrLysAspHis 268
DB 755 AGTTACAACTGGCACTACCTATGATATTTTGGGAGGCTCTGCTTGAACCCAGACAG 814
QY 269 LeuLysLysIleAsnGlyPheSerAsnAspPheTyrTPGlyIlyGluAspAspAsp 288
DB 815 TATATGAAGATGATGATGATTTCTTATATGATGATGATGATGATGATGATGATGAT 874
QY 289 LeuAlaThrArgTrhSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAla 308
DB 875 ATTGCACACAGGCTCGTCCGCGGTATGAAATCACCCCGCTCGTCCGCGTGGCG 934
QY 309 ArgTyrLysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyr 328
DB 935 CACTTAAAGATGCTCAACACAGGAGACCATGCGAACGAGAAATCCACACAGGTTT 994
QY 329 LysIleMetGlyGlnInThrLysArgArgTrhTrhArgAspGlyLysSerAsnLysTyr 348
DB 995 GACCTTCTTATCAGAGACCCAGAGATGTGGAAGACAGCGGATATACCTCGTGAACAC 1054
QY 349 LysLeuValAsnLeuGluLeuLysProLeuTyrTrhArgAlaValAlaLysLeu---Leu 367
DB 1055 AAGCTGTGCGGACAGAGCTGGAATCATCTTACCAACGAGACGTGTGATAGGGG 1114
QY 368 GluLysAspCysArgArgGluLeuArgArgAspPheProThr 381
DB 1115 GACCCGAGGCAAGAAACCCCTCAGAAATCGGGGTCCACCC 1156
RESULT 8
AK116794
LOCUS AK116794 2082 bp mRNA linear INV 30-NOV-2002
DEFINITION Clona intestinalis cDNA, clone:cleg035d14, full insert sequence.
ACCESSION AK116794
VERSION AK116794.1 GI:23589473
KEYWORDS FLI CDNA.
SOURCE Clona intestinalis
ORGANISM Clona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Clonidae; Clona.
REFERENCE
AUTHORS
1 Satou Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T.,
Sasaki, A., Hamaguchi, M., Awazu, S., Yagi, K., Sasaki, Y.,
Nakayama, A., Ishikawa, H., Inaba, K. and Satoh, N.,
A cDNA resource from the basal chordate Clona intestinalis
Genes 33 (4), 153-154 (2002)
2 (bases 1 to 2082)
Satou Y. and Satoh, N.
Direct Submission
Submitted (04-OCT-2002) Nori Satoh, Kyoto University, Department of
Zoology, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail:satoh@ascidian.zool.kyoto-u.ac.jp, Tel:81-75-753-4095,

COMMENT Fax:81-75-705-1113)
 Clona intestinalis cDNA Project (URL:
 http://ghost.zool.kyoto-u.ac.jp/index1.html).
 FEATURES Location/Qualifiers

source
 1.2082
 /organism="Clona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="cleg035d14"

ORIGIN

Alignment Scores:

Pred. No.: 5.22e-58 Length: 2082
 Score: 731.50 Matches: 148
 Percent Similarity: 60.18% Conservative: 53
 Best Local Similarity: 44.31% Mismatches: 120
 Query Match: 36.02% Indels: 13
 DB: Gaps: 4

US-10-661-430-1 (1-383) x AK116794 (1-2082)

QY 36 GUAAsnleuhtlleglserSerThrlleuilealspValalspAlaMetGUAlaVal 55
 DB 585 GAATGACAAATACCTTCTACACA-----ACATCTACAGTT 623
 QY 56 LeuGlyAsnThrlaserThrSerAspAspLeuaspThrTrpAsnSerThrPheSer 75
 DB 624 TTAGGTTCAACA-----ACAACGAAACCAACAACAACAACAATATCAACAACA 677
 QY 76 ProileserGUValanglnThrSerPheMetGUAspIleargProilleuPhePro 95
 DB 678 ACAACAACAGAGCAACAACAACAGCTGCTCCATGAT-----AAACCAACGAGAAGCTGA 731
 QY 96 AspAsnGlnThrleuGlnPheCyAsnGlnThrProProHlsleuValGlyProIlearg 115
 DB 732 GAAGAAAGTATTTTGGCCAGATGTGCTGGAACCCCTCCAGACTTACAGGGAATATAGAT 791
 QY 116 ValPheleu-----AspGluProAspPheleuThrlleuGlnlyslleTyProAspThr 133
 DB 792 GTGTCCTTTGCCACACACTGCACCGAATAATCCAACTAGAGCGGATACCCACTGCTT 851
 QY 134 HlsAlaGlyGlylyslGlyMetProlyAspCyValAlaargHlsargValAlaIle 153
 DB 852 AAAGATGCGCGTCACTCAAGCCCCCTTACGTATTAAGCTTATGATTAAGTTGCCATAGT 911
 QY 154 ValProThrArgAspArgGUAlaHlsleuArgIleMetleuHlsAsnleuHlsSerleu 173
 DB 912 ATTCTTATCGGAGACAGAGAAACATTTGGGATTTCTCCAGTACATGATCCACC 971
 QY 174 LeuAlaIysGlnGlnleuAspTyAlaIlePheIleValGlnGlnValAlaAsnGlnThr 193
 DB 972 CTNCAAGCTCAACAACCTTGATGACCATGTATGTATATACGGCTGGACTGTA 1031
 QY 194 PheAsnArgGlylyslleuMetAsnValGlyTyArgValAlaAspArgleuTyProTirp 213
 DB 1032 TTTAATGTGTCCAAAGTGAAGTGTGTTATCTGATCAATCAAAAGATCATGATTTT 1091
 QY 214 GlnCyAspPheIlePheHlsAspValAspLeuaspProGlnAspAspAspAsnleuTyThr 233
 DB 1092 CAATGTTTGTCTTCCATGATGTGATGCTTGGCTTGAAGATGACAAAGATATATCTCT 1151
 QY 234 CysProIleGlnProArgHlsMetSerValAlaIleAspTyPheAsnTylyslleuPro 253
 DB 1152 TGCCCTCAAGTCTCTCGCACCTGTACGCTGGGTGACAAAGTTAATTAACCACTGCC 1211
 QY 254 TySerAlaIlePheGlylyslleuSerAlaIleuThrlYsAspHlsleuTylyslleuAsn 273
 DB 1212 TACTCTCCATATTTGGTGGGTGATGACCTGACAAAGGAACATTTCCAAAAGTGAAC 1271
 QY 274 GlnPheSerAsnAspPheThrlleuGlylyslleuGlnAspAspAspAlaIleThrlArgThr 293
 DB 1272 GGTTACTTATTCATTTCTGGGGTTGGGTGGAGAAAGCAAGATATGTTTAACCGTGC 1331

QY 294 SerMetAlaGlyleuAspValSerArgTyProThrglnIleAlaArgTylyslMetIle 313
 DB 1332 AAGTTTCTGGCATGAATATATTCGTTACCTTATGATATATCTCAAGGATTAAGATGAT 1391
 QY 314 TyHlsSerThrlGUAlaThrlAsnProValAsnlyslCysArgTylyslMetGlyGln 333
 DB 1392 ACTCATCAAAAGAGAGAGAAAGAAACGAACCAACCCAAACGCTTCCACCAATATCCAG 1451
 QY 334 ThrTyArgArgTyThrlArgAspGlyleuSerAsnleuTylyslleuValAsnleu 353
 DB 1452 ACTTAAGACACGATGGCAACGATGCTTAAACACACTTGATATCAAGTATCAAG 1511
 QY 354 GlnleuTyProleuTyThrlArgAlaValAspLeu 367
 DB 1512 CAGAAGAACAGCTTATATCAATGTATGACGTTGACATTATG 1553

RESULT 9
 BC041742
 LOCUS
 DEFINITION
 XENOPUS LAEVIS SIMILAR TO UDP-GALACTOSE 4-EPIMERASE, COMPLETE CDS.
 ACCESSION
 BC041742
 VERSION
 BC041742.1 GI:27469876
 KEYWORDS
 MGC.
 SOURCE
 XENOPUS LAEVIS (AFRICAN CLAWED FROG)
 ORGANISM
 XENOPUS LAEVIS
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Xenopodinae; Xenopus; Xenopus.

REFERENCE
 AUTHORS
 Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.
 1 (bases 1 to 2626)
 Genetic and genomic tools for Xenopus research: The NIH Xenopus
 initiative
 Dev. Dyn. 225 (4), 384-391 (2002)
 JOURNAL
 PUBMED
 12454917
 REFERENCE
 AUTHORS
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,Y., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaby,S.J., Bosak,S.A., McKwan,P.J., Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulty,S.W., Vallejo,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bonfield,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalins,D.E., Scherf,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 PUBMED
 12477932
 REFERENCE
 AUTHORS
 Klein,S. and Strausberg,R.
 Direct Submission
 Submitted (23-DEC-2002) National Institutes of Health, Xenopus Gene
 Collection (XGC), National Institute of Child Health and Human
 Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
 20892-7510, USA
 NIH-MGC Project
 20892-7510, USA
 REMARK
 COMMENT
 Contact: XGC help desk
 Email: cgaaps-remail.nih.gov
 Tissue Procurement: Dr. Igor David
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Institute for Systems Biology
 http://www.systemsbio.org
 contact: amadan@systemsbiology.org
 Anup Madan, Jessica Fahay, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNU at: http://image.llnl.gov
 Series: IRAX Plate: 94 Row: n Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source

1..2626

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="MGC:52827 IMAGE:4884596"

/tissue_type="Embryo, stage 17/19, Xenopus"

/clone_lib="NICHD_XGC_Emb2"

/lab_host="DH10B"

/note="vector: pCMV-SPORT6"

1..2626

/gene="MGC52827"

/db_xref="GeneID:379106"

96..1592

/gene="MGC52827"

/codon_start=1

/product="MGC52827 protein"

/protein_id="AAH41742.1"

/db_xref="GI:27469877"

/db_xref="GeneID:379106"

/translation="MRVLEPSPCTLMLIGFQAFMYVFSFGSLAFIFGRSSEPS
FDYETHTVYSLSLVNSVLPSTEDTLTPCQPSRSPVNPISVTSSQDPTAEIH
KNKPLVFGGEYKPNCEKRNNTAIIIPRNEFHRLHYLHFFLQROQLHTYI
IHQAGNSTFNRAKLNVGKEALRDEMDCLFLHVDVLIPEMDFLVYCDPSPGHAS
VAMKFSYSLPYMPTFGVSALTPOYMKMGFPNRYWGMGEEDDIATRYALGSMKI
TPSPVGVHYKWKAKHGDQGNENBPHFLILRTORMKVIDMNSILNTKLARELEPL
YTNVTVDIGVDPRAKALRNRPPTPAISHYINTKTPADGQINPAPDHPNDSGIN
EAEISAEPRNKPVVAPESTTAEKLSIPMASAPSAQGEDKOISVIRBMKNNGRVK
EYQARVKKKSIVQSSKILEROVELFGAREKRPQDVGSPLPHLSAPL"

324..1130

/gene="MGC52827"

/note="Galactosyl T 2; Region: Galactosyltransferase. This is a family of galactosyltransferases from a wide range of Metazoa with three related galactosyltransferases activitys"

/db_xref="CD:DFam02709"

misc_feature

324..1130

/gene="MGC52827"

/note="Galactosyl T 2; Region: Galactosyltransferase. This is a family of galactosyltransferases from a wide range of Metazoa with three related galactosyltransferases activitys"

/db_xref="CD:DFam02709"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-661-430-1 (1-383) x BC041742 (1-2626)

9.58e-58

730.00

60.61%

43.94%

35.94%

5

Length:

Matches:

Conservative:

Mismatches:

Indels:

Gaps:

2626

145

55

120

10

5

5

5

5

5

5

5

360 TCCGTCACCTTCAGCAATCCCAACCTCGAGAGAGATTCCAAACAGAACCCCTCGTG 419

134 HislAaglygYhlsGlyMetProlysaspCyvAlaAagHlsArgValAlallele 153

420 TTCAAAGGTGGAGATACCAAGCCCAACGTGTAGAGCCCGGACCAACAGCCGATCATC 479

154 ValProlyrAaAaPArgGlyAlaHlsleuArgllemetLeuHlsAvalleuHlsSerleu 173

480 ATCCCTCACCGGACAGACAGACATCCTGAGGACCTGCTCTTATTACTGACCCCTTC 539

174 leuAlaLysGlnGlnLeuaspTyrallelperHlsValGlnGlnValAlaAaGlnThr 193

540 CTGCAGCCGACAGACCTCCTACATCAGATCTACATCATCATCAGGCGGATTAACCCAG 599

194 PheAlaLysGlyLysLeuMetAsnValGlyTyrraspValAlaSerArgLeuTyrrProThr 213

600 TTCACCGGGCCAGCTGTCTAATGTGGGTGAAGAGCTCTGGGAGATGAGGACTGG 659

214 GlnCysPheilePheHlsaspValaspLeuaspProGlnAaPArgAaAsnLeuTyrrThr 233

660 GACTCTCTGTTCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 719

234 Cys---ProileGlnProArgHlsMetSerValAlallelaspLysPheAsnTyrrLysLeu 252

720 TGTGACCCCTGAGGAGCCCAACAGCCCTGTAGGACATGAACAAATTGACATACAGCTG 779

253 ProTyrrSerAlallelPheGlyGlyLeSerAlalleuTyrrLysaspHlsleuLyslele 272

780 CCATACCTTATGTATTTTGGGGAGTCTCTGCCCTTAACCCCAACAGCATATATGAAGATG 839

273 AsnGlyPheSerAsnaspPheTrpGlyTyrrGlyGlyGlnAaPArgPLeuAlaThrArg 292

840 AATGATTTCTTAATGATATTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899

293 ThrSerMetAlaGlyLeuLysValSerArgTyrrProThrGlnlelAaArgTyrrLysMet 312

900 GTCCCTCTGGGCGGATGAATAATACCCGCTCGCTGTCGTGGGCGCATATGAAGATG 959

313 IleLysHlsSerThrGlnAlaThrAsnProValAsnLysCysArgTyrrLysleleleGly 332

960 GTCAAGACCAAGAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1019

333 GlnThrLysArgArgTyrrThrArgaspGlyLeuSerAsnLeuLysTyrrLysleuValAsn 352

1020 AGGACCCAAAGATGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1079

353 leuGlnLeuLysProLeuTyrrThrArgAlaValValaspLeu---LeuGlnLysaspCys 371

1080 AGAGAGCTTAAGAACTCTCTACACCAATGTGACCGGTGATATATGAGGAGGAGGAGGAGG 1139

372 ArgArgGlnLeuArgArgArgPheProThr 381

1140 AAGAAAGCCCTCAGAAATCGGGGTCACAC 1169

RESULT 10

BC075452 2340 bp mRNA linear VRT 24-NOV-2004

LOCUS BC075452

DEFINITION Xenopus tropicalis UDP-Gal:betaGalicnAc beta 1,4-

galactosyltransferase, polypeptide 3, mRNA (CDNA clone MGC:89236

IMAGE:7017963), complete cds.

ACCESSION BC075452

VERSION BC075452.1 GI:49522603

KEYWORDS MGC.

SOURCE Xenopus tropicalis (Silurana tropicalis)

ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Xenopus; Silurana.

1 (bases 1 to 2340)

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

REFERENCE

AUTHORS

TITLE	2 (bases 1 to 2340)
JOURNAL	Klein,S. and Gerthard,D.S.
PUBMED	Direct Submission
REFERENCE	Submitted (28-JUN-2004) National Institutes of Health, Xenopus Gene
AUTHORS	Collection (XGC), National Institute of Child Health and Human
TITLE	Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
JOURNAL	20892-7510, USA
REMARK	NIH-MGC Project
COMMENT	Contact: XGC help desk Email: cgabbs-remail.nih.gov Tissue Procurement: Rob Granger, University of Virginia cDNA Library Preparation: Open Biosystems cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Motin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhur, Parvaneh Saeedi, JR Santos, Angelique Schmech, Ursula Skalska, Diane Small, Jeff Stott, Miranda Trai, George Yang, Jacque Schein, Aaim Siddiqui, Rob Holt, Marco Marra.	
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI, at: http://image.llnl.gov Series: IRAX Plate: 171 Row: P Column: 1 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysts, Similarity but not identity to protein. Location/Qualifiers	
FEATURES	
source	1. 2340 /organism="Xenopus tropicalis" /mol_type="mRNA" /db_xref="taxon:8364" /clone="WGC:89236 IMAGE:7017963" /tissue_type="whole body, male, 10 months old, strain F6, normal" /clone_1fb="NICHD_XGC_Swd1" /lab_host="NDH10B" /note="vector: pXpress1" 1. 2340 /gene="D4gal13-3'prov" /note="synonym: MGC89236" /db_xref="GeneID:448369" 227..1291 /gene="D4gal13-3'prov" /codon_start=1 /product="UDP-gal:betaG1cNAc beta 1,4- galactosyltransferase, polypeptide 3" /protein_id="AAH75452.1" /db_xref="GI:49522604" /db_xref="GeneID:448369"
gene	
CDS	

ORIGIN

[illegible]

Db	932	CTCACCAGCATTTACATGAGGATCAAGGAGACTGCCCAACAAATTACTGGGGCGGGCG	991
Qy	284	GIyGluApsApApApLeuAlaThArqThiSerMetAlaGlyLeuLysValSerArqTyr	303
Db	992	GGAGAGGACGATGACATTTGCTTATGAGGATCCGCTACGGGGAAGAGCATGCTCGGACC	1051
Qy	304	ProthrGlnIleAlaArgTyrLysMetCylIleLysHisSerThrGlnAlaThAsnProVal	323
Db	1052	CCACTCAGCGCGCGCGATATTAATGATCTCCACGACCGGAGACTCGGGCAAGAGTg	1111
Qy	324	AsnLysCysArqTyrLysIleMetCylGlnThrLysArqArqTyrThArqAspGlyLeu	343
Db	1112	AACAGTCAAAAGATTCAGACCACTACGACAACCCGGCAACAATGAGGAGAGACGGAGTg	1171
Qy	344	SerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThArqAlaVal	363
Db	1172	AACTCGGCTGAGCTCAAACTGCTCTCCAGAACCAAGCCGACCTTATACCAATATCACT	1231
Qy	364	ValAspLeuLeuGln 368	
Db	1232	GTGGATATCGGGGAG 1246	
RESULT 11			
LOCUS	AF142671	1952 bp	mRNA linear ROD 01-JAN-2000
DEFINITION	Mus musculus beta-1,4-galactosyltransferase III mRNA, complete cds.		
ACCESSION	AF142671		
VERSION	AF142671.1	GI:6651181	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.		
AUTHORS	Lo, N.-W., Shaper, N.L. and Shaper, J.H.		
TITLE	Murine beta-1,4-galactosyltransferase family members		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1952)		
AUTHORS	Lo, N.-W., Shaper, N.L. and Shaper, J.H.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-APR-1999) Oncology, Johns Hopkins, 600 N. Wolfe St., Baltimore, MD 21287, USA		
FEATURES	Location/Qualifiers		
source	1..1952		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
CDS	235..1422		
	/note="beta4GalT-III"		
	/codon_start=1		
	/product="beta-1,4-galactosyltransferase III"		
	/protein_id="AAF22221.1"		
	/db_xref="GI:6651182"		
	/translation="MLRRLLRPCTTALLVSGQAVMMYSLGGFRSLASLPGDPPEP TFPDSHPHDVYNSLSLPPAAGAGAPPAQALPYCPERSPELVGVSVSPVSLAE IVENRPRVESGGRYRPAEGCEPRSRSTAIVPHRABNHRLILYHILHFLQOQALYAG YVHQANGFENRAKLINVGREALRDEEMCLFLHVDILPENNHLIYDDPGRPHR VAVMMKREKSLPYROYGGVSLTPPOYLKMGFEPNVEWVGGEEDDILARVLAGMG KISAPPTSGVGHYMKYKRGDKGNEENRHRPELRYRTONSWTQDMMNSTYRLLARBEIG PLYNTITADIGTDPRGPRAPSGPRYPPOSSOAFQEMLQRRPPARPGPLPANTHTAPP GSH"		
ORIGIN			
Alignment Scores:	1,94e-57	Length:	1952
Pred. No.:	725, 00	Matches:	134
Score:	64.96%	Conservative:	44
Percent Similarity:	48.91%	Mismatches:	94
Best Local Similarity:	35.70%	Indels:	2
Query Match:	9	Gaps:	2
DB:			

Qy	ProAbpsAnGlnThrLeuGlnPheCysAsnGlnThrProProHistiLeuValAlaGlyProIle	114
Db	445 CTTCCGGCTCAAGCATTGCCCTTACTGTCCAGAAAGATCAACCTCTTATGATGGAGTCCCTG	504
Qy	115 ArgVal---PheLeuAspGlnProAspPheLeuThrLeuGlnIubysIleTyProAspThr	133
Db	505 TCAGTATCTTTAGGCCAGTCCGCTCACTAGCGGAGATGTGGAGCGGAATCCCCGGGTG	564
Qy	134 HisAlaGlyTylGlnIleGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle	153
Db	565 GAATTCAGGGGGCCGGTACCGTCTCTGCAGGGGTGTGAGCCTGCTCCCAACAGCATATAT	624
Qy	154 ValProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu	173
Db	625 GTGCCCATCTGTGCCCCGGGAGCACACCTTCGCGTGTGCTTATCACTCGACACCTTTC	684
Qy	174 LeuAlaLeuSgIngluLeuAspTyrTalaIlePheIleValGluGlnValAlaAsnGlnThr	193
Db	685 CTGACAGCGCAGCAGCTGTGGTACGGCATTTATGTCAACACAGCGCTGGAAATGAAAG	744
Qy	194 PheAsnArgGlyLysLeuMetAsnValAlaGlyTyrAspValAlaSerArgLeuTyProTrp	213
Db	745 TTTAAACAGGGCAAAAGCTGTGAACGTAGGGGTGAGGAAACCCCTTCGTGATGAAGAAATG	804
Qy	214 GlnCysPheIlePheHisAspValAlaAspLeuLeuProGluAspAspArgAsnLeuTyThr	233
Db	805 GACTGCTTGTCTTACACGACGTGACCTCTCTCCAGAAAACGACATTAACCTTATGTG	864
Qy	234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyLysLeu	252
Db	865 TGGCAGCCCGGGGAGCCCGCCACGTTGCTGTTGCCATGAACAAGTTGGATACAGCCTC	924
Qy	253 ProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIle	272
Db	925 CCGTACCCCAAGTACTTTGGCGGAGTTTCAGCGCTCACTCCGACGACGTACTGAAGATG	984
Qy	273 AsnGlyPheSerAsnAspPheTrpGlyTyrTrpGlyGlyLysAspAspLeuAlaThrArg	292
Db	985 AACGGCTTCCCAACGAGTACTGGGGCTGGGGGTGGCAGGATGACGACATTTGTAACAGG	1044
Qy	293 ThrSerMetAlaGlyLeuLysValSerArgTyProThrGlnIleAlaArgTyLysMet	312
Db	1045 GTCCGCGCTGGCTGGATGGAAGATCTCTCGACCACTCACTCTGGGAGACATCAAAATG	1104
Qy	313 IleLeuHisSerThrGlnAlaThrAspProValAsnLysCysArgTyLysIleMetGly	332
Db	1105 GTGAAGACAGAGGGGATTAAGAAATAGAGAAAACCCCAAGATTGACCTCTGGTC	1164
Qy	333 GlnIleTyrAspArgTyTrpTrpArgAspGlyLeuSerAsnLeuLysTyLysLeuValAsn	352
Db	1165 CGTACCCAGAAATCTTGACACAAAGATGGAATGAACTACATTAACGTACCGACTGCTGGCA	1224
Qy	353 LeuGlnLeuLysProLeuTyTyrThrArgAlaValValAspLeu 366	
Db	1225 AGAGAGCTGGGTCTCTCTATACCAACATCACTGCAGACATC 1266	
RESULT 12		
BC013619		
LOCUS		
DEFINITION		
Mus musculus UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 3, mRNA (cDNA clone MGC:11711 IMAGE:3965561), complete cds.		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
Mus musculus (house mouse)		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Muridae; Murinae; Mus.		
REFERENCE		
1 (bases 1 to 1962) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,		

DB 1138 GTGAAGCAGAGGGGATTAAGAAATGAGAAATCCCAAGATTGACCTCTGGTC 1197
QY 333 GlnThrLysArgArgTTrpThrArgApGlyLeuSerAsnLeuTyrllysIleuValasn 352
DB 1198 GGTACCACGAATTTCTTGACACAAAGATGAATGAATCACTCAACGACGACTGCGCA 1257
QY 353 LeuGluLeuLysProLeuTyrlThrArgAlaValAlaAspLeu 366
DB 1258 AGAGAGCTGGGCTCTCTATACCAACATCACTGACGACATC 1299
RESULT 13
LOCUS BT020682
DEFINITION Bos taurus UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase,
BT020682
ACCESSION BT020682 1927 bp mRNA linear MAM 16-FEB-2005
VERSION BT020682.1 GI:59857728
KEYWORDS FLI_CDNA.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
Pecora; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 1927)
AUTHORS Smith,T.P., Grose,W.M., Preking,B.A., Roberts,A.J., Stone,R.T.,
Cassett,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chitko-McKown,C.G., Percea,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J. and Keefe,J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
JOURNAL 11282978
PUBMED 2 (bases 1 to 1927)
REFERENCE Harhay,G.P., Sonstegard,T.S., Clawson,M.L., Heaton,M.P.,
AUTHORS Keefe,J.W., Snelling,W.M., Weidmann,R.T. and Smith,T.P.L.
TITLE Sequencing and analysis of Bos taurus full-length insert cDNA
clones
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1927)
AUTHORS Harhay,G.P., Sonstegard,T.S., Clawson,M.L., Heaton,M.P.,
Keefe,J.W., Snelling,W.M., Weidmann,R.T. and Smith,T.P.L.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2005) Molecular Genetic Research Unit, USDA-ARS
U.S. Meat Animal Research Center, P.O. Box 166, Spur 18D, Clay
Center, NE 68933
COMMENT Gregory P. Harhay at harhay@email.marc.usda.gov.
FEATURES
source location/Qualifiers
1..1927
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="1BOV28F7; plate 28 row F column 7"
/tissue_type="pooled"
/clone_1b="1BOV"
/lab_host="Escherichia coli DH10B"
/note="vector: pCMV SPORT6; Site 1: NotI, Site 2: SalI;
library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
1..1927
/gene="B4GALT3"
/note="putative, similar to GeneID: 8703"
221..1411
/gene="B4GALT3"
/codon_start=1
/product="UDP-Gal:betaGlcNAc beta 1,4-
galactosyltransferase 3"
/protein_id="AA008699.1"
/db_xref="GI:59857729"
/translation="MKRLRPCTLLALVSGQLAVMYLISGGRSLALFGRQEP
AFDYPHPDIVSNLSHMPGAPVAPGGLPAPQGLPYCPSPSLVLPISVSFVSLA
EIVENPRVPGGRYPARCEPRRTALIVPRAREHHRLRLYLHAPFLQOOLAYG
IVIHQAGNGTFNRAKLINVGVEALRDEWMLFLHVDLLPNDHNLVYCDPRGR

ORIGIN
HVAAMNKGYSLPYQFGVSALTPOULNMNGFPNIEWMGEDDDIATRVLAG
MKISRPTSVGHKRVKVRGDKGNBENHRFDLIVRTONSWTQGNLSLTVQLSREL
GLITNTITADIGTDPGRGRTSSGHPHFGSSQARQENLQRRPARRPGPLTANHTAP
HGSH"
Alignment Scores:
Pred. No.: 4,51e-57 Length: 1927
Score: 721.00 Matches: 136
Percent Similarity: 64.29% Conservative: 44
Best Local Similarity: 48.57% Mismatches: 98
Query Match: 35.50% Indels: 2
DB: 4 Gaps: 2
US-10-661-430-1 (1-383) x BT020682 (1-1927)
QY 89 lIeaRgProIleuPheProAspAsnGlnThrLeuGlnPheCyAsnGlnThrProPro 108
DB 416 GTTGCCCGAGGGGCGCTCCAGCTCCCAAGCGCTGCATCTGCGCAAGATCTCCT 475
QY 109 HisLeuValAlaIleIleValProTyrlArgAspArgGluAlaHisLeuArgIleMetLeu 127
DB 476 CTCTTAGTGGGTCCCATATCCGTGTCTTTAGCCCAAGTGCATGCTGCGAGAGATTGTG 535
QY 128 LysIleTyrlProAspThrHisAlaGlyIleHisGlyMetProLysAspCyLeuAlaArg 147
DB 536 GAGAGGAATCCCGGGGTGAACCGGGGGCCGATACCCCTCGCAAGGTGACGCCCGG 595
QY 148 HisArgValAlaIleIleValProTyrlArgAspArgGluAlaHisLeuArgIleMetLeu 167
DB 596 TCCCGCAGCGATCATGTGTGCGCCACCGTCCGGGAGCACCACCTGCGCTGCTC 655
QY 168 HisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrlAlaIlePheIleValGlu 187
DB 656 TACCACCTGCACCCGCTTCCTGCAGCGCAGAGCTGTGCTTACGGGATTATGTCACAC 715
QY 188 GlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrlAspValAla 207
DB 716 CAGCTGGAAATGAAACATTTTAACAGGGCAACCTCGAATGTGCGGGTGGGAGGCC 775
QY 208 SerArgLeuTyrlProTlPrlCysPheIlePheHisAspValAspLeuProGluAsp 227
DB 776 CTGCAGATGAAGAATGGAGCTGTTGTTCTGCATGATGTGATCTCCTGCAGAGAAC 835
QY 228 AsparGlnLeuTyrlThrCys---ProIleGlnProAlaGlyMetSerValAlaIleAsp 246
DB 836 GACCACCACTGTATGTGTGGGATCCCGAGACCTCGGCACCTGTGCTGCATGAAC 895
QY 247 LysPheAsnTyrlLysLeuProTyrlSerAlaIlePheGlyIleSerAlaLeuThrLys 266
DB 896 AATTTGGATACAGCTCCCGTATCCCAAGTACTTTGGCGGGGTTCAGCGCTCACTCT 955
QY 267 AspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTlPrlGlyTyrlGlyIleAsp 286
DB 956 GACCAGTACCTGAAGATGATAGGTGCTTCCCAATGAATATCGGGGTGGGTGTAGAGAT 1015
QY 287 AsparPheLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArgTyrlProThrGln 306
DB 1016 GACGACATGTGTACCAAGGTTTCGTTGGCTGGAGGAATATCTGCCCCCACTCC 1075
QY 307 lIeaIaRgTyrlLysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCys 326
DB 1076 GTGGACACTTATAGATGTGAAGCACCGAGAGATTAAGGCGACGAGAAAATCCCAAC 1135
QY 327 ArgTyrlLysIleMetGlyGlnThrLysArgArgTlPrlThrArgAspGlyLeuSerAsnLeu 346
DB 1136 AGGTTAGACTCTTCGTGCGATACCAAGATTCCTGCACTCAAGATGATGAATCACTG 1195
QY 347 LysTyrlLysLeuValAsnLeuGlnLeuLysProLeuTyrlThrArgAlaValAlaAspLeu 366
DB 1196 ACATACAGTTGCTGTCTCGAAGCTGGCCCTCTTATACGAACATCACAGCAGACATT 1255
RESULT 14

BC087436 1329 bp mRNA linear VRT 01-FEB-2005
LOCUS BC087436
DEFINITION Xenopus laevis hypothetical LOC496041, mRNA (cDNA clone MGC:99248
IMAGE:7210136), complete cds.
ACCESSION BC087436
VERSION BC087436.1 GI:56270475
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 1329)
AUTHORS Klein S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 1329)
AUTHORS Strausberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Teshigahara, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramsen, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hult, S.W.,
Vallalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boutard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
TITLE Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 3 (bases 1 to 1329)
AUTHORS Klein, S. and Gerhard, D.S.
TITLE Direct Submission
JOURNAL Submitted (02-DEC-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
REMARK NIH-MGC Project
COMMENT Contact: XGC help desk
Email: xgcs-remail.nih.gov
Tissue Procurement: Igor B. Dawid
cDNA Library Preparation: Express Genomics
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
FEATURES
Source
1..1329
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"

/clone="MGC:99248 IMAGE:7210136"
/cise_type="Testis, adult male, wild type"
/clone_id="NICHD_XGC_Te2"
/lab_host="DH10B"
/note="Vector: pExpress1"
1..1329
/gene="LOC496041"
/db_xref="GeneID:496041"
190..1254
/gene="LOC496041"
/codon_start=1
/product="LOC496041 protein"
/protein_id="AAH87436.1"
/db_xref="GI:56270476"
/db_xref="GeneID:496041"
/translation="MPLSRIDKHNFLLLICFSQLFLILYRGAAGVFGQVFGSDP
PMDYSKIDVYTNLSVFPENPVNTQYCPASPIIVGFLVSFQPTPLKXVOQNR
LYKPGGYSFPRHCFGKRYTAVIIPHNRRARLTLVYDHPLOQLKFLVHQ
GNSFNRAKLNIQVREALKDEWCLIHVDVLPENDYNDYICDEEYKILASMD
KHHVSLPYWYTFGVSAALPPDQMYNIGFNPNSYMGSGEDDIAMIRLAGMSITPT
LSIGRYKMTSHNRDSGENSKRYDQLGNTRTRTMEDKNSLDFKLISRTAPLVYNI
TYALGDAPLVETTPAQMKLF"

ORIGIN
Alignment Scores:
Pred. No.: 3,54e-57 Length: 1329
Score: 720.00 Matches: 143
Percent Similarity: 57.46% Conservative: 61
Best Local Similarity: 40.28% Mismatches: 117
Query Match: 35.45% Indels: 34
Gaps: 5
US-10-661-430-1 (1-383) x BC087436 (1-1329)
QY 14 LeuLeuValLeuCyAlaValLeuLeuValHisAlaMetIleTyrIys-----Ile 31
DB 223 CTGCTCTGCTGCTGCTTCTCTGCTGCTTCTTCTTAACTATACGAGAGGGCGCT 282
QY 32 ProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuLeuAlaSerVala 51
DB 283 GCCGGCGCTTCCAGAGCGCTGTTGGGGATCCGAT----- 318
QY 52 MetGluAlaValLeuGlyAsnThrAlaSerThrSerAspAspLeuAspThrTPaSn 71
DB 319 -----CCGCGCTTGGGAC 330
QY 72 SerThrPheSerProIleSerGluValAsnGlnThrSerPheMetGluSerIleArgPro 91
DB 331 TACTCCAAACCATGATGATGATACACCAACCTGAGTGTGTGTGCCCAAC----- 381
QY 92 IleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuVal 111
DB 382 -----CCAGAGAAATGTAATACTCACTATGCCCCAGCAAAATCACCATACTGGTT 432
QY 112 GlyProIleArgVal---PheLeuAspGluProAspPheLeuThrLeuGluIuYrIleTyr 130
DB 433 GGTCTCTTATAGTTCATTCATTCGAAACCCCCGACCTGAAAGAGTGCAGCAAGAAAC 492
QY 131 ProAspThrHisAlaGlyIleGlyMetProIleAspCysValAlaArgHisArgVal 150
DB 493 CGATATGTGAAGCGCTGCGCTTCTTCAAGCCGACGACACTGCTTGGCGCTACAGAAAC 552
QY 151 AlaIleIleValProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnLeu 170
DB 553 GCGCTCTTATATCTTACACCGAAACAGAGAGCGCACCTCCGACCTCTTACTACTG 612
QY 171 HisSerLeuLeuAlaIysGlnGlnLeuAspTyrAlaIlePheIleValGlnValAla 190
DB 613 CACCCCTTCTTCAAGGCCAGCAGCTTATATGACCTTCAATGTCACACGAGCGCGG 672
QY 191 AsnGlnThrPheAsnArgGlyIleuMetAsnValGlyTyrAspValAlaSerArgLeu 210
DB 673 AACAGACATTCATGACGACCAAACTGCTAAACATTGGGGGTGAGGAGGCCCTPAAAC 732

[illegible]

```

        /mol_type="mRNA"
        /db_xref="taxon:9913"
        /clone="3BOV57P3; plate 57 row p column 3"
        /tissue_type="pooled"
        /clone_id="3BOV"
        /lab_host="Escherichia coli DH10B"
        /note="Vector: pCMV SPOR6; Site 1: NotI; Site 2: SalI;
        library made from pooled tissue from marrow, alveolar
        macrophage, ovary, fetal semitendonsus muscle, and fetal
        longissimus muscle."
        1..1921
        /gene="B4GALT3"
        /note="putative, similar to GeneID: 8703"
        215..1405
        /gene="B4GALT3"
        /codon_start=1
        /product="UDP-Gal:betaGalcnAac beta 1,4-
        galactosyltransferase 3"
        /protein_id="AA08933.1"
        /db_xref="GI:59858197"
        /translation="MLRLRLRPCTLLALVGSOLAVNMVYLSTGFRSLALFGEQRE
        APYSHHDVYSNLSHHPGAPVAPDGLPAHQGLPYCPKRPDLVGPISVSFSPVSLAL
        EIVERNPRVPEGGRKPRARCEPSPSRN11TVPHRRHRLHLLYTHPFLQROQLAG
        I1YHQAQNGTFRNRAKLNVGVEALDEBDEDCFLHHDVLLPENHNLVYCCDGRGP
        HVAWNAKPGYSLLPYQYFGSVSLATPDQYLKNGGFENEYWGMGEDDILATRYLGA
        MKLSRPTSGVHKMYKMGVRKDGKNEENPHRFDLLVRNOSWTOGMSLITQLSREIR
        GP1YNTLTADIGTDPDRGRTSSGPHYFSSQAFRQMLQRRPARRGPLPTANHTAP
        HGHSH"

```

ORIGIN

Alignment Scores:

Pred. No.:	5.56e-57	Length:	192
Score:	720.00	Matches:	136
Percent Similarity:	64.23%	Conservative:	44
Best Local Similarity:	48.57%	Mismatches:	98
Query Match:	35.45%	Indels:	2

US-10-661-430-1 (1-383) X BT020916 (1-1921)

Qy 89 I l e a r g P r o i l l e u P h e P r o a s p a n g l I n t h r L e u g l i n P h e C y s a n g l I n t h r P r o P r o 108
::: :::: ::::

109 HisLeuVal[Gl]VProIleAraVal---PheIeuAsnGluProAsnPhelVstHreugIu 122

Db 470 CTCTTAGTGGGTCCTATATCCGCTGCTTAAAGCCAGTGCACATCGCTGGCAGAGATTGTG 529

QY 128 LysIleTyrProAspThrHisAlaGlyGlyHisGlyMetProLysAspGlyValAlaArg 147

148 HisArgValAlaIleIleValProTyrArgAspArgLualahisLeuArgileuMetLeu 1677

Db 590 TCCCGGACAGGCATCATTTGCCCCACCGTGCCTCCGGAGACACCACTGGCCCTGTGCTC 649

[illegible]

188 GlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAla 207

Db 710 CAGGCTGGAATGGAACATTTAAcAGGGCAAGCTGTAATCGGGGCTGGGAGGCC 769

Db 770 CTGCGATGAGATGCGACTGCTTGTTCCTGCATATATGTGATCTCTCCGCCAGAGAAC 829

QY 228 AsparGAsnLeuTyrThrCys--ProIleGlnProArgHisMetSerValAlaIleAsp 246

247 lvsphaaantvrlvslleuproflvrsesralal[ephcg]wg[vl]eseralajleuthrlvls 266

COMMENT	Contact: Gregory P. Harhay at h
FEATURES	Location/Qualifiers
source	1. .1921
	/organism="Bos taurus"

247 LysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLys 266

```
Db      890  AAATTGGATACAGCCCTCCGTAATCCCACTACTTTGGCGGGGTCTCAGCGCTCACTCCT 949
Qy      267  AsPHisLeuLysLysIleAsnGlyPheSerAsnAspPheTrpGlyTyrGlyGluAsp 286
Db      950  GACCAGTACCTGAAGATGAATGGCTTCCCAATGAATATCTGGGGCTGGGGTGAAGAT 1009
Qy      287  AspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGln 306
Db      1010 GACGACATTCTCTACCAAGGTTCCGCTTGGCTGGATGAAGATATCTGTCCTCCCCCACTCC 1069
Qy      307  IleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCys 326
Db      1070 GTGGGACACTATAGATGTGAAGACACGAGAGATAGGGCAACGAGAAATCCCCAC 1129
Qy      327  ArgTyrLysIleMetGlyGlnThrLysArgArgTrpThrArgAspGlyLeuSerAsnLeu 346
Db      1130 AGGTTTGACCTCCCTGGTCCGTACCCAGAAATTCCTGACTCAAGATGGAATGAATCACTG 1189
Qy      347  LysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAlaValAspLeu 366
Db      1190 ACATACCGATTGCTGCTCGAGAGACTGGGCCCTCTTATACGAACATCACAGCAGACATT 1249
```

Search completed: January 9, 2006, 03:00:01
Job time : 6403 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 9, 2006, 00:36:56 : Search time 4931 Seconds
(without alignments)
3634.040 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSLLVLCNV.....VDLEKDCRRLRDPFPCF 383

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_plus.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10661430/runat_06012006_123135_6248/app_query.fasta_1.519
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10661430@CGN_1_1_5315@runat_06012006_123135_6248 -NCPU=6 -ICPU=3
-NO WMAP -LANG=OJURY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_hic:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_esc7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1212.5	59.7	860	7	CV125688 OSTR30043
2	1148.5	56.5	866	7	CV125689 OSTF30043
3	727.5	35.8	1590	4	CR614082 full-length
4	717	35.3	1196	11	DD035706 Homo sapi
5	717	35.3	1860	4	CR612341 full-length
6	717	35.3	1870	4	CR607487 full-length
7	717	35.3	1876	4	CR592876 full-length

8	717	35.3	1901	4	CR599499 full-length
9	717	35.3	1979	4	CR925982 Pongo pyg
10	693.5	34.1	911	7	CR444324 CR444324
11	692.5	34.1	1738	4	BC004095 Mus muscu
12	680.5	33.5	970	1	AL558425 AL558425
13	675	33.2	791	8	CR780748 AGENCOURT
14	675	33.2	803	8	CR744001 JGI AMBT1
15	670.5	33.0	796	8	CR757659 AGENCOURT
16	670.5	33.0	955	5	BO957788 AGENCOURT
17	669.5	33.0	879	6	CD514077 AGENCOURT
18	667.5	32.9	2369	4	BC029303 Mus muscu
19	664	32.7	737	8	CR744145 JGI AMBT1
20	658.5	32.4	844	1	AL866706 AL866706
21	658	32.4	687	5	BM054310 BM054310
22	657	32.3	759	8	CK457513 JGI XZG54
23	655.5	32.3	922	6	CP329240 AGENCOURT
24	651	32.1	907	5	BX488994 BX488994
25	649	32.0	893	8	CK362564 JGI XZT42
26	648.5	31.9	871	7	CK188001 EST777316
27	647	31.9	803	8	CK780638 AGENCOURT
28	646	31.8	1700	4	CR607401 full-length
29	645.5	31.8	876	5	BO929587 AGENCOURT
30	642.5	31.6	2303	4	CR749555 Homo sapi
31	638	31.4	661	3	BU617372 BU617372
32	638	31.4	1136	1	AL556185 AL556185
33	636.5	31.2	1035	10	AY410571 AY410571
34	634.5	31.2	857	8	CK442498 JGI XZG61
35	634	31.2	684	5	BM437339 BM437339
36	633	31.2	1368	4	CNS06MSO Tetraodon
37	633	31.2	2094	4	AK085368 Mus muscu
38	633	31.2	2103	4	AK052013 Mus muscu
39	633	31.2	2104	4	AK041692 Mus muscu
40	633	31.2	2104	4	AK050005 Mus muscu
41	633	31.2	2169	4	AK053603 Mus muscu
42	632.5	31.1	907	6	CF552958 AGENCOURT
43	628	30.9	622	5	BM347970 BM347970
44	623.5	30.7	1027	10	AY410573 Mus muscu
45	623	30.7	679	3	BU061791 BU061791

ALIGNMENTS

RESULT 1
CV125688/LOCUS
DEFINITION OSTR30043C08 PDONR201 Entry vector Caenorhabditis elegans cDNA,
mRNA sequence.
ACCESSION CV125688
VERSION CV125688.1 GI:51716887
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE 1 (bases 1 to 860)
Lamesch,P., Milstein,S., Hao,T., Rosenberg,J., Li,N., Segueria,R., Bosak,S., Doucette-Stamm,L., Vandenhaute,J., Hill,D.E. and Vidal,M.
Rhabditidae; Rhabditidae; Pelodactylidae; Caenorhabditis.
TITLE C. elegans ORFome Version 3.1: Increasing the coverage of ORFome
resources with improved gene predictions
JOURNAL Genome Res. (ORFome issue) (2004) In press
COMMENT Contact: Philippe Lamesch and Tong Hao
Marc Vidal Lab
DFCI
44, Binney Street, Boston, MA 02115, USA
Tel: 6176323910
Fax: 6176325739
Email: philippe_lamesch@dfci.harvard.edu
PCR Primers
FORWARD: TGCGTTTCGTCATTGGCAGT
BACKWARD: TAAACAACGCTGGAAGTCCCT.
Location/Qualifiers
1..860
/organism="Caenorhabditis elegans"

```

/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="male and hermaphrodite"
/tissue_type="all tissues"
/cell_type="all cells"
/dev_stage="all stages"
/clone_lib="pDONR201 Entry vector"

```

ORIGIN

```

Alignment Scores:
Pred. No.: 2,49e-135 Length: 860
Score: 1212.50 Matches: 248
Percent Similarity: 93.23% Conservative: 0
Best Local Similarity: 93.23% Mismatches: 9
Query Match: 59.70% Indels: 9
DB: 7 Gaps: 3

US-10-661-430-1 (1-383) x CV125688 (1-860)

QY 127 GULYGLIETYPProasp---ThriSaLa---GlyGIYHIGLY-MerProLYsAspCY 144
DB 788 GAAATATTCTATCCGAGCACCCGCCGCGTGACCATGAAATGCCCTTAAGGAT 729
QY 144 s--ValAlaArg-HisArgValAlaIle-IleVal-ProTYrArgAsp-ArgGluAlaH 162
DB 728 TGGTGTTCAGAGGCGCATCGTGTCTATTATTGTGCCCTATAGAGATTCGTGAAGCAC 669
QY 162 iSLeuArgIleMetLeuHis-AsnLeuHisSerLeuLeuAlaYsgInLeuAspTYr 181
DB 668 ATTGGAAATAAAGCTCCAAATAATTCACCTCGTGTCCCCCAACAACTATGACTAT 609
QY 182 AlaIlePheIleValGIUGInValAlaAsnGIleThrPheAsnArgGlyLYsLeuMetAsn 201
DB 608 GCATTTTCAATTGTGGAGCAAGTGCGCAATCAGCTTTATTCGGGGAACATAATGAC 549
QY 202 ValGIYTyAspValAlaSerArgLeuTYrProTYrGInCYsPheIlePheHisAspVal 221
DB 548 GTTGGATACAGACGTAGATCAACGCTCTACCATGGCAGTGCTTCATCTTTCATGATGTC 489
QY 222 AsPLeuLeuProGluAspAspArgAsnLeuTYrThrCYsProIleGInProArgHISMet 241
DB 488 GATTTCGCCCGAAGATGACCGTAACCTGTACCGTGTCAATTCACCAACGCTCATATG 429
QY 242 SerValAlaIleAspLYsPheAsnTYrLYsLeuProTYrSerIleIlePheGIYGLYIle 261
DB 428 AGGTGTCCGATCGATTAATTCATTAATTAACCTTCATATTCGGCGATCTTGGCGAATC 369
QY 262 SerAlaLeuThrLYsAspHisSLeuLYsLYsIleAsnGIYsPheSerAsnAspPheTYrGIY 281
DB 368 AGTGCACTAACAAAGATCACTGTGAAGAAATCAATGATTTTTCGATGATTTTGGGGT 309
QY 282 TTPGIYGIYGLUAspAspAspLeuAlaThrArgThrSerMetAlaGIYLeuLYsValSer 301
DB 308 TGGGGCGGAGAGGCGACGATTTGGCGACGAGAACATGATGCTGACCTGAAGATTTCA 249
QY 302 ArgTYrProThrGInIleAlaArgTYrLYsMetIleLYsHisSerThrGInAlaThrAsn 321
DB 248 AGATATTCGACAAATTTGACGATTAATAATGATTAAGCACTTCGACGAGAGCGACAAAT 189
QY 322 ProValAsnLYsCYsArgTYrLYsIleMetGIYGIInThrLYsArgAGTTPThrArgAsp 341
DB 188 CCAATTATTAATATCCCTACCAAAATATATGGCCAAAGCAAGCGCGATGGAACGCGAC 129
QY 342 GIYLeuSerAsnLeuLYsTYrLYsLeuValAsnLeuGInLeuLYsProLeuTYrThrArg 361
DB 128 GGCCTTAAGCAATCGAAGTATTAAGCTCGTAAATTCGAATTTGAAGCTCTCTCAACTCGA 69
QY 362 AlValAlaAspLeuLeuGInLYsAspCYsArgArgGInLeuAlaGArgAspPheProThr 381
DB 68 GCGGCTCGTCAATTTGCTGAAAAAGACTGCGCGGAGCTGCGAAGGACTTTTCAACG 9
QY 382 CyAsPhe 383

```

```

Db 8 TGTATT 3
RESULT 2
CV125689
LOCUS
DEFINITION
OSTR30043C08 pDONR201 Entry vector Caenorhabditis elegans cDNA,
mRNA sequence.
ACCESSION
CV125689
VERSION
CV125689.1 GI:51716888
KEYWORDS
EST.
SOURCE
Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 866)
Lamesch,P., Milstein,S., Hao,T., Rosenberg,J., Li,N., Segueria,R.,
Bosak,S., Doucette-Stamm,L., Vandenhaute,J., Hill,D.E. and Vidal,M.
C. elegans ORFeome Version 3.1: Increasing the coverage of ORFeome
resources with improved gene predictions
Genome Res. (ORFeome issue) (2004) In press
JOURNAL
Contact: Philippe Lamesch and Tong Hao
Marc Vidal Lab
DECI
44, Binney Street, Boston, MA 02115, USA
Tel: 6176323910
Fax: 6176325739
Email: philippe_lamesch@fci.harvard.edu
PCR Primers
FORWARD: TGGCTTTTCGTATTTGGCAGT
BACKWARD: TAAACACGTTGGAAGTCCCT.
Location/Qualifiers
1..866
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="male and hermaphrodite"
/tissue_type="all tissues"
/cell_type="all cells"
/dev_stage="all stages"
/clone_lib="pDONR201 Entry vector"

ORIGIN
Alignment Scores:
Pred. No.: 1.39e-127 Length: 866
Score: 1148.50 Matches: 236
Percent Similarity: 92.08% Conservative: 8
Best Local Similarity: 89.06% Mismatches: 14
Query Match: 56.55% Indels: 7
DB: 7 Gaps: 3

US-10-661-430-1 (1-383) x CV125689 (1-866)

QY 2 AlaPheArgHisLeuAlaValAlaArgLeuLYsSerLeuLeuValLeuCYsAlaValLeu 21
DB 3 GCTTTTCGTCATTTGGCAGCGCGCAACCAATCAATCGTGTCTGTACTTTGTGCGCTTCT 62
QY 22 LeuLeuValHisAlaMetIleTYrLYsIleProSerLeuTYrGInLeuLeuThrIleGIY 41
DB 63 CTATTGTTCATGCAATGATTTTAATAATTCATCGCTTTACAAAGACCTTACTATCGGC 122
QY 42 SerSerThrLeuIleAlaAspValAlaMetGIYAlaValLeuGInYAsnThrAlaSer 61
DB 123 TCTTCGACCTTATTTGCCGACGTCAACGCAATGGAAGCACTGCTCGGAAATACGCTTCC 182
QY 62 ThsSerAspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGIYAlaAsn 81
DB 183 ACTTCGATATCTACTTGATGATGATGGAATTCACAGCTTTTCAACGATTTCTGAAGTTAAT 242
QY 82 GInThsSerPheMetGluAspIleArgProIleLeuPheProAspAsnGInThrLeuGIn 101
DB 243 CAAACTAGTTTATGAGGACATTCGTCCCAATCTGTTCCTCCGACAAACGACTTTTCAA 302

```

QY 102 PheCysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro 121
| | | | |
Db 303 TTCTGTATACAGACACCTCCCACTCGTCGAGACCATCGGTATTTCTTCATGAGGCC 362
QY 122 AspPheLeuThrLeuGlnLysIleTyrProAspThrHisAlaGlyHisGlyMetPro 141
| | | | |
Db 363 GACTTCAAAATCTCGAGAAATCTATCCGACACGACGCGGTGACATGGAAATGCCT 422
QY 142 LysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAla 161
| | | | |
Db 423 AAGGATAGTGTGCAGAGGATCGTGTGTATATATGTCCTTATAGAGATGCTGAAGCA 482
QY 162 HisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyr 181
| | | | |
Db 483 CATTTGAGATTAATGCTCCACATTTGCACTGCTTCCTGCCAAACACATTTGACAT 542
QY 182 AlaIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsn 201
| | | | |
Db 543 GCAATTTTCATTTGTGAGCAGAGTGGGATTCAGACGTTTATCGCGGAAACTAATGAA 602
QY 202 ValGlyTyrAspValAlaSerArgLeuTyrProTyrGln-CysPheIlePheHisAspVal 221
| | | | |
Db 603 GTTGGATAGCAGATGACATCCGCTCTACCATGACAGGAGCTTCCTTCATGATGT 662
QY 221 L-AspLeuLeuProGlu-AspAspArgAsnLeuTyr---ThrCysProIleGlnPro-Arg 239
| | | | |
Db 663 CCGATTTACTGCCCGCAAAATGACCTTAACCTGGTACCCGCTGCTCAATTTACACACCG 722
QY 239 HisMetSerValAla---IleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePhe 258
| | | | |
Db 723 TCATATGAGGTGAGCCGATCCGAAATTTCAATTAAAAACCT---TCCCTTATTTC 779
QY 258 eglTgLy 260
| | | | |
Db 780 GGGGGGA 786
RESULT 3
CR614082 1590 bp mRNA linear HTC 21-JUL-2004
LOCUS Full-length cDNA clone CS0DJ005YK01 of T cells (Jurkat cell line)
DEFINITION Cot 10-normalized of Homo sapiens (human).
ACCESSION CR614082
VERSION CR614082.1 GI:50494889
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 1590)
AUTHORS Li W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Paradise Avenue
2 (baaes 1 to 1590)
REFERENCE 2 (bases 1 to 1590)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
COMMENT - Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by life technologies, a
division of invitrogen.
FEATURES
source
1.1590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2,98e-76 Length: 1590
Score: 727.50 Matches: 142
Percent Similarity: 62.21% Conservative: 49
Best Local Similarity: 46.25% Mismatches: 112
Query Match: 35.82% Indels: 4
DB: Gaps: 3
US-10-661-430-1 (1-383) x CR614082 (1-1590)
QY 64 AspAspLeuLeuAspThrThrAsnSerThrPheSerProIleSerGluVal---AsnGln 82
| | | | |
Db 44 GATTCGCTGAGAGTGTGAGAGGCTTTCTTCGCTTCGATGACATTTCCAAATTTT 103
QY 83 ThrSerPheMetGluAspIleArgProIleLeuPhe-ProAspAsnGlnThrLeuGlnPh 102
| | | | |
Db 104 AGATTAATGCTCACATCTCTGTCCCGAGGGGCTCTCAAGTCTCAAGTCTCCCTA 163
QY 102 eCysAsnGlnThrProProHisLeuValGlyProIleArgVal---PheLeuAspGluPr 121
| | | | |
Db 164 CTGTCCAGAGAGATCTCTCTCTTATGTGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 223
QY 121 AspPheLysThrLeuGlnLysIleTyr-ProAspThrHisAlaGlyHisGlyMetPr 141
| | | | |
Db 224 ATCACTGCGCAGAGATTTGGAGCGGAAATCCCGGGTAGAACCGAGGGCGCGGTACCGCC 283
QY 141 LysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAl 161
| | | | |
Db 284 TCACGTTGTAGACCCCGCTCCGACACGACATCAATGTGCTTACGCGCCGGAGGCA 343
QY 161 HisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyr 181
| | | | |
Db 344 CCACCTGGCGCTGCTGCTCTACACCTGACACCCCTTTCGACGGCCAGACGCTGCTTA 403
QY 181 rAlaIlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAs 201
| | | | |
Db 404 TGGCATCTATGATCATCCACACAGGCTGGAATGAACATTTAAACAGGGCAAACTGTGAA 463
QY 201 nValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspVal 221
| | | | |
Db 464 GCTTGGGGTGGAGAGGCTCTGCTGATGAAGAGTGGAGCTGCTTCTTGGACGATGT 523
QY 221 LAspLeuLeuProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHis 240
| | | | |
Db 524 GGAACCTTTCGCAAAATGACACACATCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGT 583
QY 240 sMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyG 260
| | | | |
Db 584 TGTTCCTGCTGTATGAACAAGTTTGATACAGCTCCCGTACCCCGACGATCTGGAGAG 643
QY 260 YLleSerAlaLeuThrLysAspHisLeuLysLysIleLeuGlnPheSerAsnAspPheTr 280
| | | | |
Db 644 AGTCTGACACTTACTCTCTGACACGATACCTGAAGAGTGAAGTGGCTTCCCAATGATACG 703
QY 280 pGlyTyrpGlyGlyAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysVal 300
| | | | |
Db 704 GGGCTGGGGTGTGAGATGACGACATTCCTACAGGGTGGCGCTGGGTGATGAAGT 763
QY 300 LserArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaTh 320
| | | | |
Db 764 CTCTGCGCCCCACATCTGTAGACACTATGAAGTGTGTGAAGCACAGAGAGATGAAGG 823
QY 320 rAsnProValAsnLysCysArgTyrLysIleMetClyGlnThrLysArgArgTyrThrAr 340
| | | | |
Db 824 CATATGAGAAATCCCAACAGATTTGACCTCTGTGTCGTAACCGAAATTCCTGGACGA 883
QY 340 gAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGlnLeuLysProLeuTyrTh 360

Db	884	AGATGGGATGAACCTACTGACATACCAACGATTGCTGGCTCGAAGAGCTGGGGCCCTTTATAC	943
Oy	360	ArgAlaValValAspLeu 366	
Db	944	CAACTCAGCAGCATT 962	
RESULT 4			
LOCUS	DQ035706	1196 bp	DNA linear GSS 02-JUN-2005
DEFINITION	Homo sapiens B4GALT3 gene, VIRTUAL TRANSCRIPT, partial sequence,		
ACCESSION	DQ035706		
VERSION	DQ035706.1		
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 1196)		
TITLE	Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civejlo,D., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.		
JOURNAL	A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees		
PUBMED	(er) PLOS Biol. 3 (6), E170 (2005)		
REFERENCE	15869325		
AUTHORS	2 (bases 1 to 1196)		
TITLE	Nielsen,R., Bustamante,C., Clark,A.G., Gianowski,S., Sackton,T.B., Hubisz,M.J., Fedel-Alon,A., Tanenbaum,D.M., Civejlo,D., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.		
JOURNAL	Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.		
FEATURES			
source	Location/Qualifiers		
	1..1196		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	<1..>1196		
	/gene="B4GALT3"		
	/locus_tag="HC4901"		
gene			
ORIGIN			
Alignment Scores:			
Pred. No.:	3.61e-75	length:	1196
Score:	717.00	Matches:	134
Percent Similarity:	64.60%	Conservative:	43
Best Local Similarity:	48.91%	Mismatches:	95
Query Match:	35.30%	Indels:	2
DB:	11	Gaps:	2
US-10-661-430-1 (1-383) x DQ035706 (1-1196)			
Oy	95	ProAbpAsnGlnThrLeuGlnPheCySaAsnGlnThrProProHisLeuValGlyProIle	114
Db	219	CCAGCTCCTCAAGGTCCTGCCCTACTGTCGACAAAGCATCTCTCTTACTGGGTCCTGG	278
Oy	115	ArgVal---PheLeuAspGlnProAspPheIysThrLeuGlnLysIleTyrProAspThr	133
Db	279	TCGGTGTCTTTAGCCCAAGTCGCATCATCGGACAGAGATTGGAGCGGAATCCCGGGTA	338
Oy	134	HisAlaGlyGlyVHieGlyMetProLysAspCySaValAlaArgHisArgValAlaIleIle	153
Db	339	GAACACAGGGGGCCGGAACCGCCTTGAGAGTTTGAGACCCCGCTCCGGAACAGCATCATT	398
Oy	154	ValProTyrArgAspArgGlnAlaHisLeuArgGlyMetLeuHisAsnLeuHisSerLeu	173

Db	339	GTGGCTCATGCTGCGCCGGAGACACCACCTGGCGCTGCTGTACCAACCTGACCCCTTC	458
Qy	174	LeuAlaLysGlnGlnInLeuAspTryTAlaIlePheIleValGlnInValAlaSerGlnThr	193
Db	459	TTGCAAGCGCGACGAGCTTGCTTATGAGCATCTATGTATCCACGAGCTGGAAATGCAACA	518
Qy	194	PheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTrp	213
Db	519	TTTAAACAGGGGCAAAACGTGTGAACGTTGGGGGTGGCGAAGCGCTTCGGTGAAGAGGG	578
Qy	214	GlnCysPheIlePheIleAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThr	233
Db	579	GACTGCGCTGTTCTTGACAGATGTGACTCTCTGGCAAAATATGACCAACATCTGTATGTG	638
Qy	234	Cys--ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeu	252
Db	639	TGTGACCCCCGGGGAGACCCCGCCCATGTGCGGTCTGTATGAACAAGTTGGATACAGCTTC	698
Qy	253	ProTyrSerAlaIlePheGlyIleSerAlaLeuThrLysAspHisLeuLysValle	272
Db	699	CCGTAACCCCCAGTACTTCGGAGAGCTCCAGACCTTACTCTGTACAGTACCTGAAGATG	758
Qy	273	AsnGlyPheSerAsnAspPheTrpGlyTyrGlyGlnAspAspAspLeuAlaThrArg	292
Db	759	AATGGCTTCCCAATGAATACTGGGGCTGGGTGTGATGATGACGATCTGTACCAAG	818
Qy	293	ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet	312
Db	819	GTGGCGCTCGCTGGGATGGAATCTCTCGGCCCCCATCTGTATGAGCACTATTAAGTG	878
Qy	313	IleLysHisSerThrGluAlaThrAsnProValAsnLysCysAlaGlyTyrLysIleMetGly	332
Db	879	GTGAAGCACCCAGAGAGATTAAGGCAATGAGAAATCCCAATTTGACCTCTGTGTC	938
Qy	333	GlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsn	352
Db	939	CGTACCCGAATTCCTGTGACCCAGATGGATGTAGTACTGACATACCACTTGCTGCT	998
Qy	353	LeuGlnLeuLysProLeuTyrThrArgAlaValAlaPheLeu	366
Db	999	CGAGAGCTGGGGCCTTTATACCAACATCAAGACAGCATTT	1040
RESULT 5			
CR612341			
LOCUS	CR612341	1860 bp	mRNA linear HTC 21-JUN-2004
DEFINITION	full-length cDNA clone CSDOK011W18 of Hela cells Cot 25-normalized		
ACCESSION	CR612341		
VERSION	CR612341.1	GI:50493148	
KEYWORDS	HTC; CNSLT cDNA.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
TITLE	Homnidae; Homo.		
JOURNAL	1 (bases 1 to 1860)		
REMARK	Li, W.B., Gruber, C., Jessee, J. and Polajcs, D.		
	full-length cDNA libraries and normalization		
	Unpublished		
	Contact: Feng Liang Email: fliang@life.techn.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
REFERENCE	Genoscope.		
AUTHORS	2 (bases 1 to 1860)		
TITLE	Direct Submissions		
JOURNAL	Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr - Web: www.genoscope.cns.fr)		
COMMENT	1st strand cDNA was primed with a NotI-oligo(dt) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.		

FEATURES

Location/Qualifiers
1. .1860
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CS0DK011Y148"
/tissue_type="Hela cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.:	7.04e-75	Length:	1860
Score:	717.00	Matches:	134
Percent Similarity:	64.60%	Conservative:	43
Best Local Similarity:	48.91%	Mismatches:	95
Query Match:	35.30%	Indels:	2
DB:	4	Gaps:	2

US-10-661-430-1 (1-383) x CR612341 (1-1860)

```
OY 95 ProAaPaenGlnThrLeuGlnPheCysAenGlnThrProHsiLeuValGlyProile 114
Db 421 CCAGCTCTCAAGGTCTGCTCCCTACTCCAGAACATCTCTCTTATGAGGCTCTG 480
OY 115 ArgVal---PheLeuAspGluProAspPheLysThrLeuGluLysIleTyProAspThr 133
Db 481 TCGGTGCTCTTAAAGCCAGTGCATTCAGGAGATTTGTGAGCGGAATCCCGGGA 540
OY 134 HisAlaGlyGlyHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 541 GAACCAAGGGGCGCGGTACCGCCCTGCAGGTTGTGAGCCCGCTCCGAAACAGCCATCAT 600
OY 154 ValProTyArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
Db 601 GTGCTCTCATGCGCCCGGAGACACACCTGCGCTGCTGTACCACTGCACCTGCACCTTC 660
OY 174 LeuAlaLysGlnGlnLeuAspTyArgAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
Db 661 TTGACAGCGCAGACAGCTTCTTATGGCATCTATGTCATCCACAGCGTGAATGGAACA 720
OY 194 PheAsnArgGlyLysLeuMetAsnValGlyTyArgAspValAlaSerArgLeuTyProTTP 213
Db 721 TTAAACAGGGCAAACTGTGAACGTTGGGCGAGAGCCCTGCTGATGAGAGAGTGG 780
OY 214 GlnCysPheIlePheHisAspValAspLeuProGlnAspAspArgAsnLeuTyThr 233
Db 781 GACTGCTGCTTTTGGACATGTGACCTCTTCCAGAAATGACACATCTGATATGTG 840
OY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyTyLysLeu 252
Db 841 TGTGACCCCGGGGAGACCCCGCATGTGCGTGTGATGACACAGATTGGATACAGCCCTC 900
OY 253 ProTySerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIle 272
Db 901 CCGTACCCCGCCAGTACTTCGAGAGAGTCTGACACTTACTCTGACACAGTCACTGAAGATG 960
OY 273 AsnGlyPheSerAspAspPheTTPGlyTTPGlyGlyValAspAspAspLeuAlaThrArg 292
Db 961 AATGGCTTCCCAAGTAATACTGGGCTGGGGTGGAGATGAGACACATCTCTACAGG 1020
OY 293 ThrSerMetAlaGlyLeuLysValSerArgTyProThrGlnIleAlaArgTyTyLysMet 312
Db 1021 GTGGCCCTGCTGGATGAGATCTCTGCGCCCGCCACATCTGTGAGACACATTAAGATG 1080
OY 313 IleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyTyLysIleMetGly 332
Db 1081 GTGAAGACCCGAGAGATTAAGGACATGAGAAATCCCGACAGATTTGACCTCTGCTGTC 1140
OY 333 GlnThrLysArgArgTyThrArgAspGlyLeuSerAsnLeuLysTyTyLysValAsn 352
Db 1141 CATTACCCAGAAATCTCTGAGCAGAAATGGATGAATCTCACTACATCAACAGTTGCTGCT 1200
OY 353 LeuGluLeuLysProLeuTyThrArgAlaValAspLeu 366
```

Db 1201 CGAGAGCTGGGGCTCTTTATATCCAAATCAGACAGACATT 1242

|||||
RESULT 6
CR607487 1870 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DC015YB21 of Neuroblastoma Cot
DEFINITION 25-normalized of Homo sapiens (human).
ACCESSION CR607487.1 GI:50488294
VERSION
KEYWORDS
SOURCE HTC; CNSLT_cDNA.
ORGANISM Homo sapiens (human)
Hom sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1870)
L1,W.B., Gruber,C., Jeessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1870)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqdef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers

FEATURES

source
1. .1870
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="CS0DC015YB21"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.:	7.1e-75	Length:	1870
Score:	717.00	Matches:	134
Percent Similarity:	64.60%	Conservative:	43
Best Local Similarity:	48.91%	Mismatches:	95
Query Match:	35.30%	Indels:	2
DB:	4	Gaps:	2

US-10-661-430-1 (1-383) x CR607487 (1-1870)

```
OY 95 ProAaPaenGlnThrLeuGlnPheCysAenGlnThrProHsiLeuValGlyProile 114
Db 421 CCAGCTCTCAAGGTCTGCTCCCTACTCCAGAACATCTCTCTTATGAGGCTCTG 480
OY 115 ArgVal---PheLeuAspGluProAspPheLysThrLeuGluLysIleTyProAspThr 133
Db 481 TCGGTGCTCTTAAAGCCAGTGCATTCAGGAGATTTGTGAGCGGAATCCCGGGA 540
OY 134 HisAlaGlyGlyHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 541 GAACCAAGGGGCGCGGTACCGCCCTGCAGGTTGTGAGCCCGCTCCGAAACAGCCATCAT 600
OY 154 ValProTyArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
Db 601 GTGCTCTCATGCGCCCGGAGACACACCTGCGCTGCTGTACCACTGCACCTGCCTTC 672
OY 174 LeuAlaLysGlnGlnLeuAspTyArgAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
Db 673 TTGACAGCGCAGACAGCTTCTTATGGCATCTATGTCATCCACAGGCTGGAATGGAACA 732
```

QY 194 pheasnarglylyseuemetasenva1g1ytraspvalalaserargleutyrrpctpp 213
Db 733 ttttaacagagcgaahactgtttgaactgttgaggtccgagagccctcgctgatagaagctgg 792
QY 214 g1ncysphe1lepheliaspvalasp1leuProgluaspasparganleutyrrhr 233
Db 793 gactgcctgttttgacagatgtagactcttccagaaaaatgacacacatctgtatgng 852
QY 234 cys---Pro1leuinProarghisMetSerValAla1leasplysphenleutyrrlsleu 252
Db 853 tctgacccccggagacccccccatggtccggtgcatagaacagatttgatcaccttc 912
QY 253 ProtyrSerAla1lephelglylyleSerAla1leuThryaspHisleu1ysile 272
Db 913 ccctaccacccagactcttcgagagactcgcacttaactcttaaccagactcgaagatg 972
QY 273 AsnglyPheSeraspPhe1rptg1ytrp1ygl1ygluaspasp1leu1athrarg 292
Db 973 aatggcttccccaatgaaatctggggctgggggtgtagagatgacgacatgtctacag 1032
QY 293 ThrSerMetAla1leu1ysValSerArgTyrrProThrgln1lealarg1ylysmet 312
Db 1033 gtccgcctgtcgtgagaaagatcttcgccccccacatctgtagacacttaagatg 1092
QY 313 1lelyHisSerThrgln1athrAsnProvalAsn1ysCysArgTyrrlysileMetGly 332
Db 1093 gtaaaacacccagagaaatgaagggcaatgagaaaatccccacagatttgacctctgctc 1152
QY 333 GlnThryAspArg1rptg1rptg1rptg1y1leuSerasp1leu1ysTyrrlsleuValasn 352
Db 1153 cgtaccagaaattctcgtgacgcaagatgggatgaactcactgacatctcagctgctgct 1212
QY 353 leuGlnleu1ysProleutyrrThrArgAlaValasp1leu 366
Db 1213 CGAGAGCTGGGGCTTTTATACCAACATCAGACAGACACTT 1254

RESULT 7
CR592876
LOCUS CR592876 1876 bp mRNA linear HTC 21-JUN-2004
DEFINITION full-length cDNA clone CSOD1008YC21 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR592876
VERSION CR592876.1 GI:50473683
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 1876)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/InvitrogenCorporation1600
Faraday Avenue
2 (bases 1 to 1876)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
location/qualifiers
1..1876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 7.13e-75 Length: 1876
Score: 717.00 Matches: 134
Percent Similarity: 64.60% Conservative: 43
Best Local Similarity: 48.91% Mismatches: 95
Query Match: 35.30% Indels: 2
DB: 4 Gaps: 2

US-10-661-430-1 (1-383) x CR592876 (1-1876)

QY 95 ProAspAsnGlnThryleuGlnPheCysAsnGlnThrProProHisleuValGlyPro1le 114
Db 434 ccaactcttcaaaagttgctccctactgtcccaaaacatctctctttagtgggtctgtg 493
QY 115 ArgVal---Phe1leuasp1uProaspPhe1yThryleuGln1y1leTyrrProaspThr 133
Db 494 tgggtctctttagcccaatgctcactgagagatgttgagagcgaatcccgagta 553
QY 134 HsAlaGly1yHisGlyMetPro1yAspCysValAlaArgHisArgValAla1le1le 153
Db 554 GAACAGGGGGCCGGTACCCGCTCGAGGTTGTGAAGCCCGCTCCGGAACAGCCATCAT 613
QY 154 ValProTyrrArgAspArg1y1aHisleuArg1leMetLeuHisAsnHisSerLeu 173
Db 614 gtccctcatatgtgcccggagacacactcgcctgtctcttaccactgacaccccttc 673
QY 174 leuAla1yGlnGlnleuaspTyrrAla1lePhe1leValGlnGlnValAlaAsnGlnThr 193
Db 674 ttgcacggccagagacttctttagtcactatgtcatccacagcctggaaatgacaca 733
QY 194 pheasnarglylyseuemetasenva1g1ytraspvalalaserargleutyrrpctpp 213
Db 734 ttttaacagcgaahactgtttgaactgttgaggtccgagagccctcgctgatagaagctgg 793
QY 214 GlnCysPhe1lepheliaspvalasp1leuProgluaspasparganleutyrrhr 233
Db 794 gactgcctgttttgacagatgtagactcttccagaaaaatgacacacatctgtatgng 853
QY 234 cys---Pro1leuinProarghisMetSerValAla1leasplysphenleutyrrlsleu 252
Db 854 tctgacccccggagacccccccatggtgagctgtagaacaagtttgatcaccttc 913
QY 253 ProtyrSerAla1lephelglylyleSerAla1leuThryaspHisleu1ysile 272
Db 914 ccgtacccccagactcttcgagagactcgcacttaactctgacacagatcctgaagatg 973
QY 273 AsnglyPheSeraspPhe1rptg1ytrp1ygl1ygluaspasp1leu1athrarg 292
Db 974 aatggcttccccaatgaaatctggggctgggggtgtagagatgacgacatgtctacag 1033
QY 293 ThrSerMetAla1leu1ysValSerArgTyrrProThrgln1lealarg1ylysmet 312
Db 1034 gtccgcctgtcgtgagaaagatcttcgccccccacatctgtagacacttaagatg 1093
QY 313 1lelyHisSerThrgln1athrAsnProvalAsn1ysCysArgTyrrlysileMetGly 332
Db 1094 gtaaaacacccagagaaatgaagggcaatgagaaaatccccacagatttgacctctgctc 1153
QY 333 GlnThryAspArg1rptg1rptg1rptg1y1leuSerasp1leu1ysTyrrlsleuValasn 352
Db 1154 cgtaccagaaattctcgtgacgcaagatgggatgaactcactgacatctcagctgctgct 1213
QY 353 leuGlnleu1ysProleutyrrThrArgAlaValasp1leu 366
Db 1214 CGAGAGCTGGGGCTTTTATACCAACATCAGACAGACACTT 1255

RESULT 8
CR599499

LOCUS CR599499 1901 bp mRNA linear HTC 21-JUL-2004
 DEFINITION full-length cDNA clone CSDD008Yp16 of Neuroblastoma Cot
 ACCESSION CR599499
 VERSION CR599499.1 GI:50480306
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1901)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 CONTACT : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue
 Genoscope.
 REFERENCE 2 (bases 1 to 1901)
 AUTHORS Direct Submission
 TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 FEATURES
 source
 1.1901
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSDD008Yp16"
 /tissue_type="Neuroblastoma Cot 50-normalized"
 /plasmid="pCMVSPORT_6"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.28e-75 Length: 1901
 Score: 717.00 Matches: 134
 Percent Similarity: 64.60% Conservative: 43
 Best Local Similarity: 48.91% Mismatches: 95
 Query Match: 35.30% Indels: 2
 DB: 4 Gaps: 2
 US-10-661-430-1 (1-383) x CR599499 (1-1901)
 Oy 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIle 114
 Db 442 CCAAGCTCTCAAGAGTGTGCTTACTGTCCAGAACATCTCTCTTAAAGGCTGCTG 501
 Oy 115 ArgVal---PheLeuAspGlnProAspPheLysThrLeuGlnLysLeuIleTyrProAspThr 133
 Db 502 TCGGTGTCCTTTAGCCAGTGCATCATCGGACGAGATTTGTGGAGCGAATCCCGGGA 561
 Oy 134 HisAlaGlyGlyHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
 Db 562 GAACCAAGGGGGCGGATACCGCCCTGAGGTGTGAGCCCGCTCCGGAACAGCCATCAT 621
 Oy 154 ValProTyrArgAspArgGlnAlaHisLeuArgGlyMetLeuHisAsnLeuHisSerLeu 173
 Db 622 GTGCTCATCTGCTCCCGGAGACACACTGCGCTCTGCTCTTACCACTGCACCCCTTC 681
 Oy 174 LeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
 Db 682 TTGACAGCGCAGAGCTTGTATGCACTATGTCATCCACAGCTGGAATGGAACA 741
 Oy 194 PheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyr 213
 Db 742 TTAAACAGGCAACAACTGTGAACTTTGGGCTGCGAGAGCCCTCGGTATGAAAGATGG 801

Oy 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThr 233
 Db 802 GACTCCCTGTTCTTTCACAGATGTGACCTTGTCCAGAAATGACACACATCTGTATGG 861
 Oy 224 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeu 252
 Db 862 TGTGACCCCGGGGAGCCCGGCATGTTCCGTTGTATGAACAACTTTGATACAGCCCTC 921
 Oy 253 ProTyrSerAlaIlePheGlyGlyIleSerAlaLeuTrrLysAspHisLeuLysLysIle 272
 Db 922 CCGTACCCCGCACTACTGAGAGAGTCTGACACTTACTCTGCACAGTACTCTGAAGATG 981
 Oy 273 AsnGlyPheSerAspAspPheTrrPglYtrPglYglYgluAspAspLeuAlaThrArg 292
 Db 982 AATGCTTCCCAATGAATATCTGGGGCTGGGGTGTGAGATGACATGCTTCAAGG 1041
 Oy 293 ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet 312
 Db 1042 GTGCGCTGCTGGGATGAGATCTCTCGGCCCCACATCTGTGAGACACTATAGATG 1101
 Oy 313 IleLysHisSerThrLysAlaThrAsnProValAsnLysCysArgTyrLysIleMetGly 332
 Db 1102 GTGAAGCACCGAGAGATPAGGCAATGAGAAATCCCAACAGATTTGACCTCTGCTC 1161
 Oy 333 GlnThrLysArgArgTrrPThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsn 352
 Db 1162 CGTACCCAGAAATCTCTGACGCAATGAGATGAGATCACTGACATCACAGTTGCTGGCT 1221
 Oy 353 LeuGluLeuLysProLeuTyrThrArgAlaValValAspLeu 366
 Db 1222 CGAGAGCTGGGGCTCTTTATTCACAACTACACAGACATCATT 1263
 RESULT 9
 CR925982 1979 bp mRNA linear HTC 06-DEC-2004
 LOCUS Pongo pygmaeus mRNA; cDNA DKFZp459J2220 (from clone DKFZp459J2220) .
 DEFINITION CR925982
 ACCESSION CR925982
 VERSION CR925982.1 GI:56403663
 KEYWORDS HTC.
 SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Pongo.
 REFERENCE 1 (bases 1 to 1979)
 AUTHORS Bloeker, H., Boeher, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,
 Osanger, A., Fobo, G., Han, W. and Wiemann, S.
 COMMENT The German cDNA Consortium
 Direct Submission
 Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
 Neuberg, GERMANY
 CONSRM Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 JOURNAL Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by GBF (National Research Centre for Biotechnology Ltd.,
 Braunschweig/Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 This clone (DKFZp459J2220) is available at the RZPD Deutsches
 Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
 Please contact RZPD for ordering:
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp459J2220
 Further information about the clone and the sequencing project is
 available at http://mips.gsf.de/projects/cdna/
 FEATURES
 source
 1.1979
 /organism="Pongo pygmaeus"
 /mol_type="mRNA"
 /db_xref="taxon:9600"
 /clone="DKFZp459J2220"
 /tissue_type="cortex"
 /clone_lib="459 (synonym: pcor1). Vector pSport1_Sfi; host
 DH10b; sites SfiIa + SfiIb"
 /dev_stage="adult"
 /note="UDP-Gal:beta-GlcNAc beta-1,4-galactosyltransferase

gene 3 (Homo sapiens) "
1. .1979
/gene="DKFZp459J2220"
271. .1452
/gene="DKFZp459J2220"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA129630.1"
/db_xref="GI:56403664"

/translation="MLRLLERPTCTALLVGSOLAVMMYLSIGFESLSALRPGDGP
TFDYSHPRDVTSLSHPGAPGPPAPQGLPYCPESPPLVGVSVSFSFVSLATV
ERNPVRVPGGRYPACGCEPRSRPILVPHAPRHHLRLYLHAPLQROQLVGLYV
IHQAGNGTFENRATLNVGVEALRDEMDCLPLADVDLPLNDHNLVYCDPRGRHVA
VNAKKFGSLPYPOYFGVGSALTPDDQLKNGGPNEXYMGWGGEDDILATRVLAGMKI
SRPSTVGHYKQVTKHKGDDKNGENPHRFLLVTRTNSMTQDGKNSLYOLLARLGPL
YTNITADICTDPRGPRAPSGPRYPSSQAFROMLQRRRPARPGRPPRANHTALRGS
H"

ORIGIN

Alignment Scores:

Pred. No.: 7,73e-75 Length: 1979
Score: 717.00 Matches: 134
Percent Similarity: 64.60% Conservative: 43
Best Local Similarity: 48.91% Mismatches: 95
Query Match: 35.30% Indels: 2
DB: 4 Gaps: 2

US-10-661-430-1 (1-383) x CR925982 (1-1979)

```

QY 95 ProAspAsnGlnThrLeuGlnPheCyAsnGlnThrProProHisLeuValGlyProIle 114
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 475 CCAGCTCCTCAAGGTCGCCCTACTGTCCAGAACATCTCCTCTTAGTGGGCTGTG 534

QY 115 ArgVal---PheLeuAspGlnProAspPheLeuGlnGluGlyIleTyrrProAspThr 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 535 TCGGTGCTCTTAGCCCAAGTGCATCACTGGCAGAGTTGTGAGCGGAAATCCCGGGTA 594

QY 134 HisAlaGlyGlnHisGlyMetProLyAspCyAsnValAlaArgHisArgValAlaIleIle 153
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 595 GAACACGGGGGCGCGTACCGCCCTGCAGGTTGTAGCCGCCCTCCGCAACAGCATATT 654

QY 154 ValProTyrrArgAspArgGluAlaHisLeuArgGlnMetLeuHisAsnLeuHisSerLeu 173
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 655 GTGCTCATGTCGTCGGGAGACACACTGGCGCTGCTGTACCACTGACCCCTTC 714

QY 174 LeuAlaGlnGlnGlnLeuAspTyrrAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 715 TTGACGGCCAGACGCTTGTATGCACTTATGTCAATCCACAGCTGGAATGAAACA 774

QY 194 PheAsnArgGlyLeuMetAsnValGlyTyrrAspValAlaSerArgLeuTyrrProTyr 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 775 TTTAACAGGGGCAAACTGTGAACGTTGGGGTGGAGAGGCCCTCGGTGATGAAGATGG 834

QY 214 GlnCyPheIlePheHisAspValAspLeuLeuProGlnAspAspArgAsnLeuTyrrThr 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 835 GATTGCTGTCTTCTTGATGATGTGACCTTGTGCACAAATAATGACCAATCTGTATGTG 894

QY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspHisPheAsnTyrrLeu 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 895 TGTACCCCGGGGACCCCGCATGTGCGCTGTCTGAACAAAGTTTGATACAGCCTC 954

QY 253 ProTyrrSerAlaIlePheGlyGlyIleSerAlaLeuThrLyAspHisLeuValSlsIle 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 955 CCGTACCCCGCATCTTCTGAGAGAGTCTCGCGCTTACTCTGACCACTGTAAGATG 1014

QY 273 AsnGlyPheSerAsnAspPheTyrrGlyTyrrGlyGlnAspAspAspLeuAlaThrArg 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1015 AATGGCTTCCCATGATGATCTGGGGCTGGGGTGTGATGATGATGATGATGATGATGATG 1074

QY 293 ThrsrMetAlaGlyLeuValSerArgTyrrProThrGlnIleAlaArgTyrrLeuMet 312
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1075 GTGCGCTGCTGGAGATGAATCTCTCGGCCCCCACTGTGAGACACTATTAAGATG 1134

```

QY 313 IleLyHisSerThrGlnAlaThrAsnProValAsnLyCyAsnTyrrLyIleMetGly 332
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1135 GTGAACACCGAGAGATGAAGGCAACGAGAAATCCCAAGATTTACCTCTCGTTC 1194

QY 333 GlnThrLyAsnArgTyrrThrArgAspGlyLeuSerAsnLeuLyTyrrLyLeuValAsn 352
::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1195 CGTACCCCAAAATTCCTCGAGCAGCAAGATGGATGATCACTACATACCAATTCGTGGCT 1254

QY 353 LeuGlnLeuLyProLeuTyrrThrArgAlaValAsnLeu 366
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1255 CGAGAGTTGGGCTCTTTATACCAATCAATCAGCAGACATT 1296

RESULT 10

CR444324 911 bp mRNA linear EST 19-JUN-2004
LOCUS CR444324
DEFINITION CR444324 XGC-tailbud Xenopus tropicalis cDNA clone TTBA074c18 5',
mRNA sequence.

ACCESSION CR444324
VERSION CR444324.1 GI:48969911
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

REFERENCE
AUTHORS Xenopus tropicalis (western clawed frog)
TITLE Xenopus tropicalis EST project 2001 (2004)
JOURNAL Unpublished (2004)
CONTACT: Croning MDR

```

Db      194 ATAGAGTCACGCTTCCGAGAGAACTTTACTGAGAGAGTGAGAGAAAAGAACCTTAC 253
Qy      133 ThrHisAlaGlyHisGlyMetProlysaAspCysValAlaArgHisArgValAlaIle 152
Db      254 GTGTCCAAAGAGAGGGGCTTACCAAGCCAGCTGAGCTCCATCATTAATAAAGCGCGTG 313
Qy      153 ILevalProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSer 172
Db      314 ATCATCCCTCACCAGCGGAGAGAGAGAGACCTTAATAATACCTCTGATATATCTGCACCCG 373
Qy      173 LeuLeuAlaLeuGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGln 192
Db      374 TTTCTGCACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 433
Qy      193 ThrPheAsnArgGlyTyrLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrPro 212
Db      434 ACTTTCACACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 493
Qy      213 TTPGInCysPheIlePheHisAspValAlaPheLeuLeuProGluAspAspArgAsnLeuTyr 232
Db      494 TGGAGACTGTTGTTCTACACAGATGTCAGACCTCATTCCTCCGAGAGAGAGAGAGAGAGAT 553
Qy      233 ThrCysProIleGlnProArgHisMetSerValAlaIleAspIlePheAsnTyrIleLeu 252
Db      554 ACCTGCGACAATTTCCCAACACCGCTGATCGCATGAGACAAGTTTGATTAACAATTTG 613
Qy      253 ProTyrSerAlaIlePheGlyIleSerAlaLeuThrIleAspHisIleLeuIleValIle 272
Db      614 CCTTCAAAATCTTACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 673
Qy      273 AsnGlyPheSerAsnAspPheTyrPglTyrPglTyrGlyIleAspAspAspLeuAlaThrArg 292
Db      674 AACGGTTCCTCCCAACAACACTACTGCGATGCGGCGGAGAGAGATATATATAGCATCAAG 733
Qy      293 ThrSerMetAlaGlyLeuIleValSerArgTyrProThrGlnIleAlaArgTyrIleMet 312
Db      734 GTGCGACTTACGGGAGATGATCATCTCTCCCTTATCATGAGACCGGAGAGATATAAATG 793
Qy      313 ILeuValHisSerThrGluAlaThrAsnProValAsnLeuCysArgTyrIleMetGly 332
Db      794 ATCAAGACAGGCGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 853
Qy      333 GlnThrIleArgArgTyrThrArgAspGlyLeuSerAsnLeuIleTyrIleLeuVal 351
Db      854 AAGACCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 910

```

RESULT 11
BC004095
LOCUS
DEFINITION Mus musculus UDP-gal-4-epimerase beta 1, 4-galactosyltransferase, polypeptide 3, mRNA (cDNA clone IMAGE:3594143), containing frame-shift errors.
ACCESSION BC004095
VERSION BC004095.1
KEYWORDS GI:14708685
SOURCE HTC.
ORGANISM Mus musculus (house mouse)
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridea; Muridae; Murinae; Mus.
1 (bases 1 to 1738)
Strusberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, P.S., Wagner, L., Shemmen, C.M., Schlier, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Scheffer, C.F., Bhac, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carinci, P., Prange, C., Raha, S., Logunov, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,

Fahney, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butcherfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Schnerich, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1738)
Strusberg, R.
Direct Submission
Submitted (28-FEB-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LINL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdopaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IMAGE Plate: 11 Row: 1 Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
This clone has the following problem: frame shifted.
Location/Qualifiers
1. 1738
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3594143"
/tissue_type="mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_11b="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Alignment Scores:
Pred. No.: 5,856-72 Length: 1738
Score: 692.50 Matches: 133
Percent Similarity: 64.60% Conservative: 44
Best Local Similarity: 48.54% Mismatches: 95
Query Match: 34.10% Indels: 3
DB: 4 Gaps: 2
US-10-661-430-1 (1-383) x BC004095 (1-1738)
Qy 95 ProAspArgnGlnThrLeuGlnPheCysAsnGlnThrProGlnIleValGlyProIle 114
Db 264 CTTCCGGCTCAAGAGATGCGCTTACTGTCACGAADAATACCTTTTACTGAGGCTCTG 323
Qy 115 ArgVal---PheLeuAspGlnProAspPheIleGlnGlyIleTyrProAspThr 133
Db 324 TAGATATCTTTAGCCCGGCGGTCATCAGACAGATGTCGAGCGGAATCCCGGGTG 383
Qy 134 HisAlaGlyGlyHisGlyMetProIleAspCysValAlaArgHisArgValAlaIleIle 153
Db 384 GAATCAGGGGCGGCTACCGTCTCAGAGGAGTGAAGCTCGCTCCGAGACCATTAAT 443

QY 154 ValProTyrArgAspArgGluAlaHisIleuValArgIleMetLeuHisIleAsnLeuHisSerLeu 173
Db 444 GTGCCCATCCAGTCCCGGAGGACCACTTCGGCTGGCTCTATATCACTCCGACCCCTTC 503
QY 174 LeuAlaIleGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaIleGlnThr 193
Db 504 CTCACACGCGACGAGCTTGGCTACGGCATTTATGTCTATCCACAGCGCTGGAAATGGAACG 563
QY 194 PheAsnArgGlyIleuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyr 213
Db 564 TTTAACAGGCAAAAGCTGCTGAACGTAAGGAGTGAAGGAGACCTTCCTGATGAAGATGG 623
QY 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThr 233
Db 624 GACTGCTGTGTTCTTACACGACGTGACCTCTTCCAGAAAACGACCTGATATGTG 683
QY 234 Cys---ProIleGlnProAlaGlnIleMetSerValAlaIleAspIlePheAsnTyrIleu 252
Db 684 TGGACCCCGGAGGACCCCGCACGTTGCTGTGCGCATGAAAGATTGA--TACAGCCTC 741
QY 253 ProTyrSerAlaIlePheGlnGlyIleSerAlaIleuThrIleAspHisIleuIleuIle 272
Db 742 CCGTACCCCGACGATCTTGGCGAGTTTCAGCGCTCACTCCGACGACGTAAGATG 801
QY 273 AsnGlyPheSerAsnAspPheTyrIleTyrGlyGluAspAspAspLeuAlaThrArg 292
Db 802 AACGGCTTCCCAACGAGTACGTGGGCTGGGAGTGGAGATGACGATTTGCTACAGG 861
QY 293 ThrSerMetAlaGlyLeuIleuValSerArgTyrProThrGlnIleAlaArgTyrIleuMet 312
Db 862 GTCGCGCTGGCTGGAGTGAAGATCTCTGACCACTCACTCTGGGACACTTAAAGATG 921
QY 313 IleuHisSerThrGlnAlaIleThrAsnProValAsnIleuIleuIleuIleuIleu 332
Db 922 GTGAAGACAGAGGAGGATTAAGAAATGAGAAATCCCAACATTTGACCTCTGTC 981
QY 333 GlnThrIleAspArgTyrIleThrArgAspGlyLeuSerAsnLeuIleuIleuIleuIleu 352
Db 982 CGTACCCAGAAATCTTGTGACCAAGATGAAATGAAATGAAATGAAATGAAATGAAATG 1041
QY 353 LeuGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 366
Db 1042 AGAAGCTGGGCTCTCTATACCAACATCACTGACAGATC 1083

RESULT 12
AL558425 970 bp mRNA linear EST 02-APR-2004
DEFINITION Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION AL558425 Homo sapiens cDNA clone CS0DU005YK01 5-PRIME, mRNA sequence.
VERSION AL558425.3 GI:46183823
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 970)
Li, W. B., Gruber, C., Jeesee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31280224.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
345.r
For more information about this cluster, see

FEATURES
source
http://www.genoscope.cns.fr/cdna?e=CS0DU005AF01QPL&c=325.r.
Location/Qualifiers
1..970
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DU005YK01"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.: 6,84e-71 Length: 970
Score: 680.50 Matches: 139
Percent Similarity: 62.00% Conservative: 47
Best Local Similarity: 46.33% Mismatches: 110
Query Match: 33.51% Indels: 5
DB: 1 Gaps: 3

US-10-661-430-1 (1-383) x AL558425 (1-970)
QY 64 AspAspLeuAspPheThrTrpAsnSerThrPheSerProIleSerGluVal---Aengln 82
Db 44 GATTCCTGCTGGAGAGCTGGATGAGAGCTTTCTGCTCCCTGTGACATTTCCATTTT 103
QY 83 ThrSerPheMetGluAspIleArgProIleLeuPhe-ProAspAsnGlnThrIleuGln 102
Db 104 AGATAATGCTCAGATCTGCTCCCGACAGGGGCTCTCCAGCTCTCAAGTGTGCCCTTA 163
QY 102 eCysAsnGlnThrProProHisIleuValGlyProIleArgVal---PheLeuAspGluTrp 121
Db 164 CTGTCCAGAAACGATCTCTCTTGTAGTGGCTGTGTGGTGTCTCTTAAGCCCAAGTCC 223
QY 121 oAspPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 141
Db 224 ATCATCTGGACAGATTTGGAGCGGAATCCCGGAGTGAACCAAGGAGCGGACCGCC 283
QY 141 oLysAspCysValAlaArgHisIleuValAlaIleIleValProTyrArgAspArgGluAl 161
Db 284 TGCAGGTTTGAGACCCCGCTCCGACAGCATCATTTGCTCATGCTGACCGGAGCA 343
QY 161 aHisIleuArgIleMetLeuHisIleuHisIleuHisIleuHisIleuHisIleuHisIleu 181
Db 344 CCACTGCGCTGCTCTCTTACCACTGCAACCCCTTCTTGACGCGCAGAGCTTCTTA 403
QY 181 fAlaIlePheIleValGluGlnValAlaIleGlnThrPheAsnArgGlyIleuMetAs 201
Db 404 TGGCATCTATGTATCATCACAGGCTGGAATGAAATTTTAACAGGCAAAATCTTTGAA 463
QY 201 nValIleGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspVa 221
Db 464 CGTTGGGGGTGCGAGAGGCTCGCTGATGAAAGTGGAGTGGCTGTCTTTCACAGATGT 523
QY 221 lAspLeuLeuProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHis 240
Db 524 GGAACCTTCCCGAAGAAATGACCAACATCTGTATGTGTGACCCCGGAGACCCGCCA 583
QY 240 sMetSerValAlaIleAspIlePheAsnTyrIleuProTyrSerAlaIlePheGlyI 260
Db 584 TGTTCGCTGTGCTATGAACAAGTTTGATACAGCTCCCGTACCCCAAGTACTTCGGAGG 643
QY 260 yIleSerAlaIleuThrIleuAspHisIleuIleuIleuIleuIleuIleuIleuIleuIleu 280
Db 644 AGCTGACGACCTTACTCTGACCAAGTACCTGAAAGATGAATGGCTCCCAATGAATACG 703
QY 280 pGlyTyrGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyIleuIleuVal 300

Db 704 GGGCTGGGTGGTGAAGATGACACATTCTACACAGGTCGCTGGCTGGAGTGAAGAT 763
Qy 300 lseratgtyrprothglnlealargtyrlysmetllelyshiserthglualath 320
Db 764 CTCTGGCCCCCAGCTCTGTAGACATATAGAGACCTGAGACCGGAGATGAGG 823
Qy 320 rAnpProValAenLysCyAargTyrlYsileMetGlyInThLyAArgTrpThrAr 340
Db 824 CATAGGAAATCCACAGATTGA-CTTCKKCCCTTACCCAGATTCTCGACGCA 882
Qy 340 gAapGlyLeuSerAsnLeuLysTyrlYsLeuValAenLysGluLeuLysProLeuTy 359
Db 883 AGATGGATGAATCTACTACATACAGTTGGTGGCTGAGACTGGGGCTCTTAT 940
RESULT 13
CX780748 791 bp mRNA linear EST 28-JAN-2005
LOCUS AGENCOURT 42763460 NIH_XGC_tropincl Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7795167 5', mRNA sequence.
CX780748
VERSION CX780748.1 GI:58297541
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 791)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Bruce Blumberg
cDNA Library Preparation: B. Blumberg
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM15945 row: d column: 13
High quality sequence start: 13
High quality sequence stop: 714.
Location/Qualifiers
1. 791
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7795167"
/tissue_type="Intestine"
/dev_stage="Adult"
/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"
/clone_lib="NIH_XGC_tropincl"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: XhoI; The
library was prepared from 5 ug of poly A+ RNA by oligo-dT
priming
(5'-GAGAGAGAGAGAGAGAGACTAGTCTGAGTTTCTTTTCTTTT-3')
and StrataScript reverse transcriptase. After ligation of
EcoRI adapters (5'-AATTCGACAGAGG-3') followed by kinasing
adapters and by XhoI digestion, the cDNA was size selected
by chromatography on Sepharose CL-2B columns and fractions
containing cDNAs larger than 1000 bp were ligated into
EcoRI/XhoI-digested pCS107. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Michelle Tabb and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."

ORIGIN

Alignment Scores:
Pred. No.: 2,326-70 Length: 791
Score: 675.00 Matches: 117
Percent Similarity: 68.70% Conservative: 52
Best Local Similarity: 47.56% Mismatches: 75
Query Match: 33.23% Indels: 2
DB: 8 Gaps: 1
US-10-661-430-1 (1-383) x CX780748 (1-791)
Qy 123 PheLysThrLeuGluLysIeTyrrProAspThr-----HisAlaGlyGlyNHisGlyMet 140
Db 9 TTGGACACCATTCATTCATTCGCGCAGGCTGTCGCCGAAGAGGGCGCTTACAG 68
Qy 141 ProLysAspCyValAlaArgHisArgValAlaIleIleValProTyrrAspArgGlu 160
Db 69 CCACCCGACTGGAGCTCCATCTATAAACGGCGGATGATCTCCACCGCGGAGAG 128
Qy 161 AlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnLeuAsp 180
Db 129 CAGCACCTTAAATACCTGCTGATATATCTGCACCCGTTTCTGCAGCGGACGACGTAAC 188
Qy 181 TyraIaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuTc 200
Db 189 TACGGGATATACATCATCATCATCAGCGGGATATTCTTCAACCGGCTAAGTTGCTG 248
Qy 201 AsnValGlyTyrrAspValAlaSerArgLeuTyrrProTrpGlnCyPheIlePheHisAsp 220
Db 249 AACGTGGCTTCAGAGAGCGCATGAAGACGAGATGGAGACTGTTGTTCTTACACGAT 308
Qy 221 ValAspLeuLeuProGluAspAspArgAsnLeuTyrrCyrrProIleInProArgHis 240
Db 309 GTGCAGCTTATCTCCGGAAGAGATGCAATATCTTACTGCGAAGAAATTTCCAAAGC 368
Qy 241 MetSerValAlaIleAspLysPheAsnTyrlYsLeuProTyrrSerAlaIlePheGly 260
Db 369 GCCTGATCGCCATGACAAAGTTTGATGAATTAATTCCTTCAATCACTTACCTCGGGGA 428
Qy 261 lIeSerAlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAspHeTrp 280
Db 429 GTGTAGCCCTTTCCTCCAGAACATATCATGAAGATGAACGGCTTCCCAACAATCTG 488
Qy 281 GlyTrpGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysVal 300
Db 489 GGATGGCGCGCGGAAGATGATATATAGCATCATCAGAGTTCGACCTTACGCGCATGATCATC 548
Qy 301 SerArgTyrrProthglnlealargtyrlysmetllelyshiserthglualathr 320
Db 549 TCTCGCCCTTCTATCCAGACCGGAGGTATAAATGATCAAGCAGCGCATGACAAAGGCG 608
Qy 321 AanProValAenLysCyAargTyrlYsileMetGlyInThLyAArgTrpThrAr 340
Db 609 AATGAGCAAAATCCCAAAAGGTTTAACATGCTGACGAAGACCGCGGACGCGGCGAG 668
Qy 341 AapGlyLeuSerAsnLeuLysTyrlYsLeuValAenLysGluLeuLysProLeuTyrr 360
Db 669 GACGGGATGAATCCCTGACAGTACTGCTCTCTCTCAAGAGATTGACAGCGCTTTACACC 728
Qy 361 ArgAlaValAlaAspLeu 366
Db 729 AACATCATCTGATGACATC 746
RESULT 14
CX744001 803 bp mRNA linear EST 24-JAN-2005
LOCUS JGI ANBT1739.fwd NIH_XGC_tropbas6 Xenopus tropicalis cDNA clone
DEFINITION IMAGE:7514094 5', mRNA sequence.
CX744001
VERSION CX744001.1 GI:58070284
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 803)
Richardson, P., Lucas, S., Rokhsar, D., Dettler, J. C., Ng, D. C.,
Brostein, P. and Lindquist, E. A.
DOE Joint Genome Institute Xenopus tropicalis EST project
Unpublished (2004)
Other ESTs: JGI ANBT1739.rev
Contact: Lindquist, E. A., Richardson, P.
DOE Joint Genome Institute
2800 Mitchell Drive, Walnut Creek, CA 94598, USA
Tel: 925 296 5600
Fax: 925 296 5710
Email: cdna@jgi-psf.org
Tissue Procurement: Robert M. Grainger
CDNA Library Preparation: Bruce Blumberg Laboratory, University of
California, Irvine
DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
Clone Distribution: I.M.A.G.E. Consortium/LLNL:
<http://image.llnl.gov>
Naming Conventions: EST name is generated by the concatenation of
the JGI Clone id and the direction of sequencing. The suffix '.fwd'
indicates a forward sequencing read of the insert. It does not
necessarily reflect the orientation of the insert.
Plate: ANBT 0017 row: e column: 4
High quality sequence stop: 738.
Location/Qualifiers

1. 803
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7514094"
/tissue_type="whole embryo"
/dev_stage="Gastrula (st. 10-13)"
/lab_host="TOP10 E. coli"
/clone_1lb="NH XGC, tropGaa6"
/note="Vector: PCS22+; Site 1: PstI; Site 2: XhoI. The
library was prepared from 5 ug of poly A+ RNA by oligo-dT
priming (V(T)20CTCGAGAGAGAGAG). After ligation of PstI
adapters and by XhoI and PstI digestion, the cDNA was size
selected by chromatography on Sepharose CL-4B columns and
fractions containing cDNAs larger than 500 bp were ligated
into PstI/XhoI-digested PCS22+. Reference for library
construction: Current Genomics 4, 635-644. Library
constructed by Jisong Peng and Bruce Blumberg (Dept of
Developmental and Cell Biology, University of California,
Irvine)."

FEATURES
source

Alignment Scores:
Pred. No.: 2,38e-70 Length: 803
Score: 675.00 Matches: 115
Percent Similarity: 69.92% Conservative: 50
Best Local Similarity: 48.73% Mismatches: 71
Query Match: 33.23% Indels: 0
Gaps: 0

US-10-661-430-1 (1-383) x CX744001 (1-803)

ORIGIN

131 ProAaPThrHisAlaGlyGlyHisGlyMetProLysAspCyValAlaArgHisArgVal 150
12 CCTACAGTGTCCAAAGAGGCGCTACAAAGCCACCGACTCGAGTCCACTCATTAACG 71
151 AAlaIleIleValPProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnLeu 170
72 GCCGTGATCATCCCTTCACCGGCGAGGAGAGACACCTTAATTCCTGCTGTATTATCTG 131
171 HisSerLeuLeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGlnGlnValAla 190
132 CACCCGTTTCTGACGCGGACCACTGAACATACCGGATATACATCATTCATCAGCGGGT 191
191 AsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeu 210

Db 192 AATTTCATCTTCACCGCGCTAAGTTCGAACGTCGCTTCACAGGAGCCCATGAAAGAC 251
Qy 211 TTTTTCATCTTCACCGCGCTAAGTTCGAACGTCGCTTCACAGGAGCCCATGAAAGAC 230
Db 252 GAGGATTCGACGCTGTTTTCACCGAGATGTCGATTCATTCGCCAAGACATTCGCAT 311
Qy 231 LeuTyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyr 250
Db 312 ATCTATACCTGGACAAATTCCTCCCAAGACGCGCTGATGCCATGACAAAGTTTGATAC 371
Qy 251 LysLeuProTyrSerAlaIlePheGlyIleSerAlaLeuThrLysAspHisIleuLys 270
Db 372 AAATTCCTTACAAATTCCTACTTCGCGGAGAGTGTACGCCCTTTCGCCAACAATACATG 431
Qy 271 LysIleAsnGlyPheSerAsnAspPheTyrGlyTyrGlyGlyLysAspAspLeuAla 290
Db 432 AAGATGACGCGCTTCGCCAACAATTCCTGCGGAGTGGCGGAGAGATGATATATAGGC 491
Qy 291 ThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyr 310
Db 492 ATGAGGTTGACACTTATGCGGATGATCATCTCGCCCTTATCCAGACGCGAGATAT 551
Qy 311 LysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIle 330
Db 552 AAATGATGACAGCAGCGCATGACAAAGGAGCAATGACAAATCCCAAAAGTTTAAACATG 611
Qy 331 MetGlyGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLysTyrLysLeu 350
Db 612 CTACAGAGACCGCGGAGAGTGGCGGAGAGAGGAGATGAACTCGCTGCATCTACTCTG 671
Qy 351 ValAsnLeuGlnLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
Db 672 CTCTCAAGAGATTGACGCGCTTTACACCAATCATCATGTGAACATC 719

RESULT 15
CX757659 796 bp mRNA linear EST 24-JAN-2005
LOCUS
DEFINITION
AGENCOUNT 40973844 NIH MGC 281 Homo sapiens CDNA clone
IMAGE:7782009 3', mRNA sequence.
CX757659 GI:58054315
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 796)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: Meri Firpo
CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM15939 row: p column: 07
High quality sequence start: 16
High quality sequence stop: 792.
Location/Qualifiers

FEATURES
source

1. 796
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7782009"

/tissue_type="pluripotent cell line derived from
blastocyst inner cell mass"
/lab_host="DHIOB Tona"
/clone_lib="NIH_MGC_281"
/note="Organ: Blastocyst; Vector: pExpress-1; Site: 1;
Score; Site 2: Not; RNA obtained from pluripotent cell
line derived from blastocyst inner cell mass (cell line
HSF-6, NIH Registry designation UC06. Positive for OCT4
expression by rtPCR, positive for SSEA-3, SSEA-4,
Tra-1-81, Tra-1-60 by immunofluorescence. Negative for
SSEA-1 by immunofluorescence Passage 62. CDNA was primed
using oligo-dT primer:
5'-pGACTAGTCTAGATCGCAGCGCGCCGCT(25-3' and cloned into
the Score/Noti sites of pExpress-1. Size-selection >1.25
kb resulted in an average insert size of 2.0 kb. This
primary library is normalized (non-normalized primary
library is NIH_MGC_280) and was constructed by Express
Genomics (Frederick, MD). Note: this is a Mammalian Gene
Collection library."

ORIGIN

Alignment Scores:
Pred. No.: 8.22e-70 Length: 796
Score: 670.50 Matches: 123
Percent Similarity: 66.12% Conservative: 39
Best Local Similarity: 50.20% Mismatches: 82
Query Match: 33.01% Indels: 1
DB: 8 Gaps: 1

US-10-661-430-1 (1-383) x CX757659 (1-796)

Qy 117 PheLeuAspGluProAspPheLysThrLeuGluLysIleTyrProAspThrHisAlaGly 136
Db 50 TTtAGCCCACTGCGCATCTGCGCAGAGATTGTGAGCGGAATCCCGGTAGAACCGAGG 109
Qy 137 GlnHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIleValProTyr 156
Db 110 GCGCGGTACCGCGCTGCAAGTTGTGAGCCCGCTCCGAAACAGCATCATTTGGCTCAT 169
Qy 157 ArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLys 176
Db 170 CGTGGCCCGGAGACACCTGCGCGCTGCTTACACCTGACCCCTTCTTGCAGCGC 229
Qy 177 GlnGlnLeuAspTyrAlaIlePheIleValGlnGlnValAlaAsnGlnInthrPheAsnArg 196
Db 230 CAGCACTTGTCTTATGCGCATCTATGTCATCCACAGCTGGAATGGAACATTTAACAGG 289
Qy 197 GlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPhe 216
Db 290 GCAGAACTGTGAGCTTGGGGTGGAGAGGCCCTGCGTGAAGAGAGTGGAGACTGCTCG 349
Qy 217 IlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCys--Pro 235
Db 350 TTCTTGACAGATGTGACCTCTTGCAGAAATGACCAACATCTGTATGTGTGACCCC 409
Qy 236 IleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSer 255
Db 410 CCGGGACCCCGGCATGTGCGCTGTGTAACAAGTTGAAACAGCTCCCGTACCCC 469
Qy 256 AlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPhe 275
Db 470 CAGTACTTGGAGAGAGTCTCAGACACTTACTCTCGACCACTGACCTGAAGATGATGCTTC 529
Qy 276 SerAsnAspPheTyrGlyTyrGlyGlyLysAspAspAspLeuAlaThrArgThrSerMet 295
Db 530 CCCAATGAATACTGGGGCTGGGGTGTGAGATGACGACATGCTACCGAGGTGCGCTCG 589
Qy 296 AlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHis 315
Db 590 GCTGGATGAAGATCTCTGAGCCCGCCACATCTGTAGACACTATAGAGATGGAGAGCAC 649
Qy 316 SerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIleMetGlyInthrLys 335
Db 335

Db 650 CGAGGAGATAAGGCAATGAGAAAATCCACAGATTTCATCTGCTCGTACCCAG 709
Qy 336 ArgArgTPTThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGlnLeu 355
Db 710 AATTCTGAGCGAAGATGGAGATGAATCACTGACATTCAGATTGCTGCTGAGAGCTG 769
Qy 356 LysProLeuTyrThr 360
Db 770 GGGCTCTTTATACC 784

Search completed: January 9, 2006, 04:21:39
Job time : 4950 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 9, 2006, 00:54:55 ; Search time 236 Seconds
(without alignments)
2884.772 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSLLVLCV.....VDLEKDCRRLRDPFPCF 383

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-O=/cg2_1/USPTO.spool/US10661430/runat_06012006_123136_6274/app_query.fasta_1.519
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=humana4.0.cdi
-LIST=45 -DLOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10661430.@CGN_1_193.@runat_06012006_123136_6274 -NCPU=6 -ICPU=3
-NO_MMAPP -LANG=ODDERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA: *
1: /cg2_6/ptodata/1/ina/1 COMB.seq: *
2: /cg2_6/ptodata/1/ina/5 COMB.seq: *
3: /cg2_6/ptodata/1/ina/6A COMB.seq: *
4: /cg2_6/ptodata/1/ina/6B COMB.seq: *
5: /cg2_6/ptodata/1/ina/H COMB.seq: *
6: /cg2_6/ptodata/1/ina/PTUS COMB.seq: *
7: /cg2_6/ptodata/1/ina/PP COMB.seq: *
8: /cg2_6/ptodata/1/ina/RE COMB.seq: *
9: /cg2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	717	35.3	1938	3	US-09-673-395A-140
2	717	35.3	1970	3	US-09-949-016-2875
3	704	34.7	1189	3	US-09-270-767-11472
4	668.5	32.9	2304	2	US-08-446-777-7
5	666.5	32.8	2304	2	US-08-446-777-5
6	666	32.8	1265	2	US-08-446-777-1
7	658	32.4	1116	3	US-09-118-464-33
8	658	32.4	1116	3	US-10-132-652A-33
9	658	32.4	1119	3	US-09-949-016-2195

10	658	32.4	1119	3	US-09-949-016-2196	Sequence 2196, Ap
11	658	32.4	2027	3	US-09-118-464-1	Sequence 1, Appl
12	658	32.4	2027	3	US-10-132-652A-1	Sequence 1, Appl
13	646	31.8	1023	3	US-09-118-464-34	Sequence 34, Appl
14	646	31.8	1023	3	US-10-132-652A-34	Sequence 34, Appl
15	642.5	31.6	1964	3	US-09-991-181-235	Sequence 235, Ap
16	642.5	31.6	1964	3	US-09-990-444-235	Sequence 235, Ap
17	642.5	31.6	1964	3	US-09-997-333-235	Sequence 235, Ap
18	642.5	31.6	1964	3	US-09-992-598-235	Sequence 235, Ap
19	642.5	31.6	2151	3	US-09-949-016-1167	Sequence 1167, Ap
20	642.5	31.6	2167	3	US-09-949-016-541	Sequence 541, Ap
21	642.5	31.6	2280	2	US-09-055-097-4	Sequence 4, Appl
22	642.5	31.6	2280	3	US-09-338-933-108	Sequence 4, Appl
23	583	28.7	2965	3	US-09-949-016-4375	Sequence 4375, Ap
24	583	28.7	4646	3	US-09-949-016-680	Sequence 680, Ap
25	577	28.4	3830	3	US-09-949-016-679	Sequence 679, Ap
26	576	28.4	3832	3	US-09-949-016-2359	Sequence 2359, Ap
27	510.5	25.1	494	3	US-09-270-767-27061	Sequence 27061, A
28	446	22.0	541	3	US-09-404-879A-108	Sequence 108, Ap
29	446	22.0	541	3	US-09-338-933-108	Sequence 108, Ap
30	446	22.0	541	3	US-09-215-681-108	Sequence 108, Ap
31	446	22.0	541	3	US-09-216-003A-108	Sequence 108, Ap
32	446	22.0	541	3	US-09-667-857-108	Sequence 108, Ap
33	446	22.0	541	3	US-10-198-053-108	Sequence 108, Ap
34	446	22.0	541	3	US-09-827-271-108	Sequence 108, Ap
35	327.5	16.1	13290	3	US-09-949-016-13937	Sequence 13937, A
36	327.5	16.1	13290	3	US-09-949-016-13938	Sequence 13938, A
37	314.5	15.5	500	3	US-09-270-767-1933	Sequence 1933, Ap
38	314.5	15.5	500	3	US-09-270-767-17215	Sequence 17215, A
39	313	15.4	10383	3	US-09-949-016-14617	Sequence 14617, A
40	272.5	13.4	1524	3	US-10-012-231A-16	Sequence 16, Appl
41	272.5	13.4	1524	3	US-10-012-231A-16	Sequence 16, Appl
42	272.5	13.4	1524	3	US-10-006-768A-16	Sequence 16, Appl
43	272.5	13.4	1524	3	US-10-015-671A-16	Sequence 16, Appl
44	272.5	13.4	1524	3	US-10-015-393A-16	Sequence 16, Appl
45	272.5	13.4	1524	3	US-10-011-833A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-673-395A-140
; Sequence 140, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673,395A
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 140
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-673-395A-140
Alignment Scores:
Pred. No.: 1.68e-82
Score: 717.00
Percent Similarity: 64.60%
Best Local Similarity: 48.91%
Query Match: 35.30%
DB: 3
Gaps: 2
US-10-661-430-1 (1-383) x US-09-673-395A-140 (1-1938)
Length: 1938
Matches: 134
Conservative: 43
Mismatches: 95
Indels: 2

Qy	95	ProArbAnGlnThrLeuGlnPheCysAsnGlnThrProPheHisLeuValGlyProIle	114
Db	432	CGAGTCTCCAGAGTCTGCCCTTACGTCCAGAACATCTCTCTTTAGTGGCTCTGTG	491
Qy	115	ArgVal---PheLeuAspGluProAspPheLysThrLeuGluLysIleTyrProAspThr	133
Db	492	TGGGTGCTTTAGGCCAGTGCATTCAGAGATGTGTGAGCGGAATCCCCGGGTA	551
Qy	134	HisAlaGlyGlnHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle	153
Db	552	GAACCAAGGAGGCGGTATCCGCCCTGCAGGTGTGAGACCCCGCTCCGAACAGCAATATT	611
Qy	154	ValProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu	173
Db	612	GTGCTCATTCGTGGCCCGGAGACACACCTGGCCCTGTCTCTACCACTGCACCCCTTC	671
Qy	174	LeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThr	193
Db	672	TTGGCAGCCCAAGACACTGTGCTTATGGCACTTATGTCATCCACAGGCTGGAATGAAACA	731
Qy	194	PheAsnATGGGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyr	213
Db	732	TTTAAACAAGGCAAACTGTGAACTGTGGGGTGCAGAGAGCCCTGGCTGATGAAGTGG	791
Qy	214	GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThr	233
Db	792	GACTCCCTGTTCTTCACCATGTGGACCTTGGCCGAAGAAATGACCAACAATCTGTATGTG	851
Qy	234	Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeu	252
Db	852	TGTGACCCCCGGGGAACCCCGCAATGTGGCTGTGCTATGAACAAGTTGGATACAGCTTC	911
Qy	253	ProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIle	272
Db	912	CCGTACCCCACTACTCTGGAGAGAGTCTCAGACCTTACTCTGACCAAGTACCTGAAGAG	971
Qy	273	AsnGlyPheSerAsnAspPheTrrpGlyTrrpGlyGlyGluAspAspAspLeuAlaThrArg	292
Db	972	AATGCTCTTCCCAATGAATATCTGGGGCTGGGGTGGAGAGATGACACAACTGTCTACAG	1031
Qy	293	ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet	312
Db	1032	GTGGCCCTGTGCTGGAGATGAAGATCTTCGGCCCCCAACATCTGTAGACCATTTAAGATG	1091
Qy	313	IleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIleMetGly	332
Db	1092	GTGAAGACCCGAGAGATGAAGGGCATGTAGGAAATCCCAACAGATTTTGACCTCTGGTGC	1151
Qy	333	GlnThrLysArgArgTrrpThrArgAspGlyLysSerAsnLeuLysTyrLysLeuValAsn	352
Db	1152	CGTACCCAGAAATTCCTGGACGCAAGATGGGATGAATCACTGACATACATCCAGTTGTGGCT	1211
Qy	353	LeuGlnLeuLysProLeuTyrThrArgAlaValAlaAspLeu	366
Db	1212	CGAGAGCTGGGGCTCTTTATACCAACATCACAGACAGACTT	1253
RESULT 2			
US-09-949-016-2875			
; Sequence 2875, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001037			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-11			
; PRIOR APPLICATION NUMBER: 60/241,755			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 60/237,768			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR APPLICATION NUMBER: 60/231,498			
; PRIOR FILING DATE: 2000-09-08			

NUMBER OF SEQ ID NOS: 207012			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 2875			
LENGTH: 1970			
TYPE: DNA			
ORGANISM: Human			
US-09-949-016-2875			
Alignment Scores:			
Pred. No.:	1.73e-82	Length:	1970
Score:	717.00	Matches:	134
Percent Similarity:	64.60%	Conservative:	43
Best Local Similarity:	48.91%	Mismatches:	95
Query Match:	35.30%	Indels:	2
DB:	3	Gaps:	2
US-10-661-430-1 (1-383) x US-09-949-016-2875 (1-1970)			
QY	95	ProAspAnGInThrLeuGInPheCyAsnGInThrProPronHisLeuValGlyProIle	114
DB	815	CCACGCTCCTCAAGGTCGCTCCCTCACTGTCACGAAGCATCTCTCTTAGTGGGTCCTGTG	874
QY	115	ArgVal---PheLeuAspGInProAspPheLeuThrLeuGInuysIleTyProAspThr	133
DB	875	TCGGATGCTCTTACCCCACTGCCATCACTGGCAGAGATGTGGAGCGGAATCCCGGCTA	934
QY	134	HisIaGlyGInHisGlyMetProLysAspCyValAlaArgHisArgValAlaIleIle	153
DB	935	GAACCAAGGGGGCGGGTACCGCCCTGCGAGGTTGTAGCGCCGCTCCGAAACGCATCATTT	994
QY	154	ValProTyTrArgAspArgGInuAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu	173
DB	995	GTGGCTCATCTGTCGGGAGACACACCTGGCGCTGCTCTACACCTTCACCCCTTC	1054
QY	174	LeuAlaLysGInGInLeuAspTyTrAlaIlePheIleValGInGInValAlaAsnGInThr	193
DB	1055	TTTGCAGCGCCACGACGCTTGCTTATGCGATCTATGTCATCCACGACGCTGGAAATGAAACA	1114
QY	194	PheAsnArgGlyLysLeuMetAsnValGlyTyTrAspValAlaSerArgLeuTyTrProTrp	213
DB	1115	TTTAAACAGGGCAAAACCTGTGAACGTTGGGGGTGGAGAGGCCCTGGATGAAGAGCTGG	1174
QY	214	GInCyPheIlePheHisAspValAspLeuLeuProGInuAspAspArgAsnLeuTyTrThr	233
DB	1175	GACTGCTGTTCTTTCGACGATGTGACCTCTTGCCAGAAATGACACAAATCTGATATGTG	1233
QY	234	Cys---ProIleGInProArgHisIleMetSerValAlaIleAspLysPheAsnTyTrLeuLeu	252
DB	1235	TGTGACCCCGGGAGCCCGCCCATGTTGCCGTGTAGAACAAATTGGATACAGCTTC	1294
QY	253	ProTyTrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleIle	272
DB	1295	CCGTAACCCCGCATCTTCGAGAGAGCTCACACACTTACTCTCTGCACAGTACCTGAAGTGG	1354
QY	273	AsnGlyPheSerAsnAspPheTrpGlyTyTrpGlyGInuAspAspLeuAlaThrArg	292
DB	1355	AATGGCTTCCCATGATTACTGGGGCGCTGGGTGTGAGATGACGATTCGTCACAGG	1411
QY	293	ThrSerMetAlaGlyLeuLysValSerArgTyTrProThrGInuAlaArgTyTrLysMet	312
DB	1415	GTGGCGCTGCTGGGATGAAATGATCTCTGGCCCCCACATCTTGAGGACACTATTAAGATG	1474
QY	313	IleLysHisSerThrGInuAlaThrAsnProValAsnLysCyAspGlyTyTrIleMetCly	332
DB	1475	GTGAAGCACCCAGAGATTAAGGCGCATATGAGAAATCCCAACAGATTGACCTCGGCTC	1533
QY	333	GInThrLysArgArgTyTrThrArgAspGlyLeuSerAsnLeuLysTyTrLeuValAsn	352
DB	1535	CGTACCCAGAAATCTCTGGAGCCAGATGTGAATCACTGACATACCAAGTTGGCTGCT	1594
QY	353	LeuGInuLeuLysProLeuTyTrThrArgAlaValValaAspLeu 366	
DB	1595	CGAGAGCTGGGGCTCTTTATACCAAACTCAACAGACGATTT 1636	

```
RESULT 3
US-09-270-767-11472
; Sequence 11472, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 11472
; LENGTH: 1189
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-11472

Alignment Scores:
Pred. No.: 3 64e-81 Length: 1189
Score: 704.00 Matches: 144
Percent Similarity: 61.49% Conservative: 46
Best Local Similarity: 46.60% Mismatches: 91
Query Match: 34.66% Indels: 28
DB: 3 Gaps: 7

US-10-661-430-1 (1-383) x US-09-270-767-11472 (1-1189)
QY 71 AanserThPhe---SerProIleSerGluValAsnGlnThrSerPheMetGluAspIle 89
DB 261 AACTCTACCATTAAGAACTGATGTCACAGGCACTTACTTCCATTCACAGACTTA 340
QY 90 ArgProIleLeu-----PheProAspAsnGln----- 98
DB 341 ACGCGCTCTCTGCGGCAAGAAATTTTCCCGCCGAGACAGAAATCCATCCGCC 400
QY 99 ThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeu 118
DB 401 CTCCTTCCCACTGACCTGATCCCGATCCCGTATGATGAGACCCATC----- 448
QY 119 ArgGluProAspPheLeuThrLeuGluValIle-----TyrPro 131
DB 449 ---AGCCCAAC---ACGACACTGGAGTCACTGAGCGTATTGAGCGGACCTTGGACCT 502
QY 132 AspThrHisAlaGlyGlyHisLeuMetProLysAspCysValAlaArgHisArgValAla 151
DB 503 CTTTTCGCCGCTGTCGCGCCCTTCGAGCCTGAACCTGCAATGCCACGATCAGCGGCT 562
QY 152 IleIleValProLysArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHis 171
DB 563 ATTTGTTGTCCTCTCGCATGATAGCGCCATCTTACTTCTTCTGCGCAACATCCAC 622
QY 172 SerLeuLeuAlaLysGlnLeuAspArgValAlaIlePheIleValGluGlnValAlaAsn 191
DB 623 CCATTTCATATAGAGAGCGCATCCGCTATCCATTTTCATTTGTAGAGACCAACGGG 682
QY 192 GlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyr 211
DB 683 AAGCCCTTCATATCGGCTGCGCATGATGAACTTTGTTATTTGAGAGCCCTTAAGCTGTC 742
QY 212 ProTTrpGlnCysPheIlePheHisAspValAspLeuLeuProGluLysAspArgAsnLeu 231
DB 743 CAGTGGAGATGTTTATATATCCAGATGTCGATCTTGCCTTGGACGACGCAATCTT 802
QY 232 TyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLys 251
DB 803 TACAACGTGCCAGTCAGCGGACGACATGTCACTATAGACGCGTGAACCTTCAAG 862
QY 252 LeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLys 271
DB 863 TTGCCTTATCATCAATATTTGAGAGTGTTCGCAATGACGCGTGAACACTTTCAGGCC 922
```

```
QY 272 IleAsnGlyPheSerAsnAspPheTrpGlyTyrProGlyGluAspAspLeuAlaThr 291
DB 923 GTAAATGGCTTCTCAAACTCGTTCTTGGCTGGCGCGCGAGATGACATGTCACAAC 982
QY 292 ArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLys 311
DB 983 AGTTGGAAGCAGCGCCAACTTATTCATCAAGATCCGTCACATGACCGCGTACAG 1042
QY 312 MetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIleMet 331
DB 1043 ATGCTGAAGCATCAGAGAGAAAGCGCCATCTTAAG-----CGCTATGAATAACTTA 1093
QY 332 GlyGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuVal 351
DB 1094 CAGATGGCATGAGAGAAATAGAACAGATGGATGCACTGATTAAGTATTTCTATAC 1153
QY 352 AsnLeuGluLeuLysProLeuTyrThr 360
DB 1154 AGCATCAAACAATTTCCAACTTTCAC 1180

RESULT 4
US-08-446-777-7
; Sequence 7, Application US/08446777
; Patent No. 5641668
; GENERAL INFORMATION:
; APPLICANT: Berger, Eric G.
; APPLICANT: Watson, Manfred
; APPLICANT: Ivanov, Svetoslav X.
; TITLE OF INVENTION: Proteins having glycosyltransferase
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: CIBA-GEIGY Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,777
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03194
; FILING DATE: 15 NOV 93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 92810924.8
; FILING DATE: 27 NOV 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferrari, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19361/A/BB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2304 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; STRAIN: E. coli DH5alpha
; IMMEDIATE SOURCE:
; CLONE: YEPGSTB
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2301
```

OTHER INFORMATION: /product=
OTHER INFORMATION: "galactosyltransferase-sialyltransferase hybrid
; OTHER INFORMATION: protein"
US-08-446-777-7

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
4,66e-76	2304	140	74	132	49	6
Score:	668.50					
Percent Similarity:	54.18%					
Best Local Similarity:	35.44%					
Query Match:	32.91%					

US-10-661-430-1 (1-383) x US-08-446-777-7 (1-2304)

QY 14 LeuLeuValLeuValAlaValLeuLeuValAlaValMetLeuTyr----- 29
DB 76 CTCGTGCGCGCTGCGTTCGACCTTGCGTCACTTCCTGTTAACTACCTGCTGCGCGC 135
QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeu 46
DB 136 GACCTAGCGCGCTGCGCCCACTG-----GTGGAAGTCTCCACACCGCTG 180
QY 47 AlaAspValAlaAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
DB 181 CAGGCGCGCTCGAAGAGTGCCTGCGCATCGGAGAGTCTCGGAGAGTCCGAGACCGA 240
QY 61 -----SerThrSer 63
DB 241 GGGGCGCGCGCGCTCTCTCTAGCGCTCTCCAGCGCGCGCGCTGCGCTCC 300
QY 64 AspAspLeuLeuAspThrTyrPheAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
DB 301 AGCCAGCTGCTGATTTGCGCCCTGCGCGCTAGCAACTGACCTCGTC----- 351
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
DB 352 -----CCAGTGCCTCCACACACACACCGACCTGCTGCGCGCTGC 390
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro---Asp 122
DB 391 CCGAGGAGTCCCGCTGCTGCGCGCTGCGCATGCTGATTAAGTTTAAACATGCTGAGAC 450
QY 123 PheLeuThrLeuGluLysIleTyrProAspThrHisIleGlyGlyHisGlyMetProLys 142
DB 451 CTGAGCTCGTGGAAAGACAAACCAATGTGAAGATGGCGCGCGCTATGCCCCAGG 510
QY 143 AspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
DB 511 GACTGCTCTCTCTCACAAGTGGCATCATCTCCATTCGCAACCGGACGAGACAC 570
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrAla 182
DB 571 CTCAGAGTACGGCTATATTTATTTGACCCACCTCTGAGCGCGCGCTGAGTATGAC 630
QY 183 IlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAspVal 202
DB 631 ATCATGTATCAACACGCGCGGAGACATATATTCATCTGTATGCTCTCAAGTTT 690
QY 203 GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAsp 222
DB 691 GGCCTTCAAGAGACCTTGAAGGACTATGACTACCTGCTTGTGTAGTGAAGTGGAC 750
QY 223 LeuLeuProGluAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
DB 751 CTCATTCATTAATGAACATTAAGCGGACAGGTTTTCACAGCGACCATTTCC 810
QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyIleSer 262
DB 811 GTTGCAATGATTAAGTTGATTCAGCTTACCTTATGTTCAGTATTTTGAAGGCTCTCT 870
QY 263 AlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAspPheTyrGlyTyr 282

DB 871 GCTTAAGTAAACAACAGTTTCTAACATCAATGATTTCTTAATATTATGGGCTGG 930
QY 283 GLYGLYGLUASPASPASPLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
DB 931 GAGGAGAGATGATGACATTTTAAACATTAAGTTTAAAGGATGCTTATCTATCTCC 990
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
DB 991 CCAATGCTGTGTGCGGAGGTGCGCATGATCCGCACTCAAGAGACAAAGAAAATGAA 1050
QY 323 ValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTyrThrArgAspGly 342
DB 1051 CCCAATCTCAGAGTTTGAACCAATTCGACACACAAAGAGACAATGCTCTCTGATGGT 1110
QY 343 LeuSerAsnLeuTyrTyrLysLeuValAsnLeuGlyLeuLysProLeuTyrThrArgAla 362
DB 1111 TTGAATCTCACTACCTTACAGGTGCTGAGTGTACAGAGTATACCATTTGTATACCCAAATC 1170
QY 363 ValValAspLeuLeuGluLysAspCysArgArgGluLeuArgArg 377
DB 1171 ACAGTGACATC-----GGACACGAGGTAGGATCGTCTGA 1206

RESULT 5

US-08-446-777-5
; Sequence 5, Application US/08446777
; Patent No. 5641668
GENERAL INFORMATION:
APPLICANT: Berger, Eric G.
APPLICANT: Iwanow, Svetoslav X.
TITLE OF INVENTION: Proteins having glycosyltransferase
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 520 White Plains Road
CITY: Tarrytown
STATE: NY
ZIP: 10591-9005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,777
FILING DATE: May 26, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/03194
FILING DATE: 15 NOV 93
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EPO 92810924.8
FILING DATE: 27 NOV 92
ATTORNEY/AGENT INFORMATION:
NAME: Ferraro, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 4-19361/A/BE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-3318
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2304 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
STRAIN: E. coli DH5alpha
IMMEDIATE SOURCE:
CLONE: YEPGSTA
FEATURE:

NAME/KEY: CDS
 LOCATION: 1..2301
 OTHER INFORMATION: /product=
 OTHER INFORMATION: "galactosyltransferase-sialyltransferase hybrid
 OTHER INFORMATION: protein"
 US-08-446-777-5

Alignment Scores:

Score:	8.48e-76	Length:	2304
Percent Similarity:	666.50	Matches:	140
Best Local Similarity:	54.188	Conservative:	74
Query Match:	35.448	Mismatches:	132
	32.828	Indels:	49
		Gaps:	6

US-10-661-430-1 (1-383) x US-08-446-777-5 (1-2304)

QY 14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleTyr----- 29
 DB 76 CTGTCGCGCGCTGCGGCTGACCTTGCGCTGACCTGTTACTGCTGCGCGC 135
 QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIle 46
 DB 136 GACCTGAGCGCGCTGCGCCCACTG-----GTCGAGTCTCCACACCGCTG 180
 QY 47 AlaAspValAlaAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
 DB 181 CAGGCGCGCTCGAACAAGTCCCGCGCATCGGCGAGTCTCTCGGGAGACTCCGACCGGA 240
 QY 61 -----SerThrSer 63
 DB 241 GGGGCGCGCGCGCTCTCTGAGCGCGCTCTCCACGCGCGCGCGCGCGCGCATCC 300
 QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
 DB 301 AGCCAGTGTGATTCGTGCGCGCGCGCGCTGAGCACTTGAACCTCGGCTC----- 351
 QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
 DB 352 -----CAGTGGCCCGACACCGCACCGCACCTGCTGCCCGCGCTCC 390
 QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
 DB 391 CCGAGAGACTCCCGCTGTCGCGCGCGCGCGCGCATGATTGATTAACTGCTGTCGAC 450
 QY 123 PheIysThrLeuGlnLysIleTyrProAspThrHisAlaGlyIleHisGlyMetProIys 142
 DB 451 CTGAGAGCTGTGCAAAAGCAGAACCCCAATGTGAAGATGGCGCGCGCTATGCCCCACAG 510
 QY 143 AspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
 DB 511 GACTGTGCTCTCTCTCAAGGTGGCATCATCATTCATCCAGAACCGCGGACGACGC 570
 QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrAla 182
 DB 571 CTCAGTACTGCTATATTATTATTGACCCAGTCTCTGACAGCGCAGAGAGTGAACATAAGC 630
 QY 183 IlePheIleValGlnGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetCysVal 202
 DB 631 ACTCATGTATCATCAACGCGGAGACACTATATTCATATCGCTAAGCTCTCATATGTT 690
 QY 203 GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAsp 222
 DB 691 GCGTTTCAAGAACCTTGAAGACATATGACTACCTGCTTTGTGTGTATGTAGCTGAGAC 750
 QY 223 LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
 DB 751 CTCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 810
 QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyIleSer 262
 DB 811 GTTGCAATGATTAAGTTGATTCAGCTTATGTTAGTTAGTTAGTTAGTTAGTTAGTTAGT 870

QY 263 AlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAspPheTrpGlyTyr 282
 DB 871 GCTCTAAGTAAACAACACTTTCTAACCATCAATGATTTCCATTAATATTATGGGCGTGG 930
 QY 283 GlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
 DB 931 GGAGGAGAAGATGATGATGATTTTAAACAGATTAGTTTAAAGGATGCTATATCTCGC 990
 QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
 DB 991 CCAAAATGCTGTGCTGGAGGCTGTCATGATCCCGCTCAAGACAAAGAAATGAA 1050
 QY 323 ValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTyrThrArgAspGly 342
 DB 1051 CCCATCTCTCAGAGGTTTGACCGAATTGCACACACAAAGAGACAAATGCTCTGATGCT 1110
 QY 343 LeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAla 362
 DB 1111 TTGAATCACTCACTTACAGAGTCTGATGATGATGATGATGATGATGATGATGATGATG 1170
 QY 363 ValValAspLeuLeuGlnLysAspCysArgArgGluLeuArgArg 377
 DB 1171 ACAATGACATC-----GGACACGAGCTGGATCCGTCGA 1206

RESULT 6

US-08-446-777-1
 Sequence 1, Application US/08446777
 Patent No. 5681668

GENERAL INFORMATION:

APPLICANT: Berger, Eric G.
 APPLICANT: Matzele, Manfred
 APPLICANT: Iwanow, Sveroslav X.
 TITLE OF INVENTION: Proteins having glycosyltransferase
 TITLE OF INVENTION: activity
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSER: CIBA-GEIGY Corporation
 STREET: 520 White Plains Road
 CITY: Tarrytown
 STATE: NY

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/446.777
 FILING DATE: May 26, 1995
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/EP93/03194
 FILING DATE: 15 NOV 93
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: EPO 92810924.8
 FILING DATE: 27 NOV 92

ATTORNEY/AGENT INFORMATION:

NAME: Ferrari, Gregory D.
 REGISTRATION NUMBER: 36,134
 REFERENCE/DOCKET NUMBER: 4-19361/A/BE
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908) 277-3318
 TELEFAX: (908) 277-4306

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1265 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

MOLECULE TYPE: CDNA

ORIGINAL SOURCE:

STRAIN: E. coli DH5alpha

IMMEDIATE SOURCE:

```

; CLONE: P4AD113
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1200
; OTHER INFORMATION: /product= "full-length
; OTHER INFORMATION: galactosyltransferase"
; US-08-446-777-1

Alignment Scores:
Score: 3 59e-76 Length: 1265
Percent Similarity: 666.00 Matches: 137
Best Local Similarity: 54.69% Conservative: 73
Query Match: 35.68% Mismatches: 128
DB: 32.79% Indels: 46
Gaps: 5

US-10-661-430-1 (1-383) x US-08-446-777-1 (1-1265)
QY 14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleYr----- 29
Db 82 CTGCTGGCGCGCTGCGCTGTCGACCTTGCGCTGACCTCGTTTACATGCTGCGCGC 141
QY 30 -----LysIleProSerLeuYrGluAsnLeuThrIleGlySerSerThrIle 46
Db 142 GACCTGAGCGCGCTGCGCCCACTG-----GTGGAGTCTCCACACCGCTG 186
QY 47 AlaAspValAlaAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
Db 187 CAGGCGCGCTGCAACAGTCCCGCGCCATCGGACATCTCCGGGAGCTCCGACCGGA 246
QY 61 -----SerThrSer 63
Db 247 GGGGCGCGCGCGCTCTCTCTAGCGCGCTCTCCAGCGCGCGCGGCTGCGACTCC 306
QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGlyValAsnGlnThr 83
Db 307 AGCCCACTCGTGAATTTGGCCCTCGCGCCGCTAGCACTGACTCGTC----- 357
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
Db 358 -----CCAGTCCCGCCACACCGCACCTGCTGCGCGCTGC 396
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
Db 397 CTGAGAGAGTCCCGCTGCTGTGTGGCCCATGCTGATTGATTAACTGCTGTGAC 456
QY 123 PheIleThrLeuGluYsIleYrProAspThrHisAlaGlyGlnYsGlyMetProLys 142
Db 457 CTGAGCTCGTGGCAAGACAGAACCCCAATGTGAAGTGGCGCGCTATGCCCGCAGG 516
QY 143 AspCysValAlaArgHisArgValAlaIleIleValProYrArgAspArgGluAlaHis 162
Db 517 GACTGCTCTCTCTCTCAAGGTGGCCATCATCTTCATTCCTCCGCAACCGCAGAGAC 576
QY 163 LeuArgIleMetCysIleAsnLeuHisSerLeuLeuAlaYsGlnGlnLeuAspTyrAla 182
Db 577 CTCAAGTACTGGCTATATATTATTTGACACCGACTCTGAGCGCGCAGCTGAGTATG 636
QY 183 IlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyYsLeuMetAsnVal 202
Db 637 ATCATATTTTCAACACGCGCGGAGACATATATATCTGCTAAGCTCTCTCAATGTT 696
QY 203 GlyTyrAspValAlaSerArgLeuYrProTrpGlnCysPheIlePheHisAspValAsp 222
Db 697 GGCCTTCAAGAAAGCTTGAAGACTATGACTACCTGCTTTGTTTGAAGTGTGAC 756
QY 223 LeuLeuProGluAspAspArgAsnLeuYrThrCysProIleGlnProArgHisMetSer 242
Db 757 CTGATTCGAATGAGATGACATATGCGTATGCGTATTTTTCACAGCCACGACATTTCC 816
QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyYsIleSer 262
Db 817 GTTGCAATGATGAATGATTGATTCAGCTTACTTATGTTCAAGTATTTTGGAGTGTCT 876

```

```

QY 263 AlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAspPheTrpGlyTrp 282
Db 877 GCTCTAAGTAAACAACAGTTTCTTAACCATCATGATTTCTTAATATATATGGGCTGG 936
QY 283 GlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
Db 937 GGAGGAGAAAGATGATGACATTTTAAACAGTATGATTTTAAAGGCAATGCTTATCTGC 996
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
Db 997 CCAATGCTGTGTGTCGGAGGCTGTGTCATGATCCGCACTCAAGACACAAAGAAATGAA 1056
QY 323 ValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTrpThrArgAspGly 342
Db 1057 CCCAATCTCAGAGGTTTGACCGAATTCACACACAAAGAACATGCTCTGTGATGT 1116
QY 343 LeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAla 362
Db 1117 TTGAACCTCACTACCTPACCGAGTGTGATGTACAGATATACCATTTGATACCAATTC 1176
QY 363 ValValAspLeu 366
Db 1177 ACAGTGACATC 1188

RESULT 7
US-09-118-464-33
; Sequence 33, Application US/09118464A
; Patent No. 6558934
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Bennett, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-glucosamine
; TITLE OF INVENTION: Beta-1,4-Galactosyltransferase, Beta4gal-T2
; FILE REFERENCE: 4305/08521
; CURRENT APPLICATION NUMBER: US/09/118,464A
; CURRENT FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-118-464-33

Alignment Scores:
Pred. No.: 3 2e-75 Length: 1116
Score: 658.00 Matches: 139
Percent Similarity: 65.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
DB: 3 Gaps: 6

US-10-661-430-1 (1-383) x US-09-118-464-33 (1-1116)
QY 7 AlaValAlaArgLeu-LysSerLeuLeuValLeuCysAlaValLeuLeuValHisAl 26
Db 24 GCTGAGAGCGGTGCAAGGCTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 83
QY 26 aMetIleYrLysIleProSerLeuYrGluAsnLeuThrIleGlySerSer---ThrIle 45
Db 84 CATCTCTTAC--TTGACGCTTACGCCCGACGCTGCGCTTCTTCAAGCGCTTCAAGTGC 140
QY 45 uIleAlaAspValAlaAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAsp 65
Db 141 CCGAGCGCTCCCGCATGCTCTCCACCGACTGCTAGCAGCAGCAGCAGCAGCAACTG 200
QY 65 pLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGlyValAsnGlnThrSerph 85
Db 201 CTCGCGGCCCAAGCCACCGCTCTAGCTCCGGGCTCCGAGGTCCGCCAGTGCC----- 255
QY 85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnGln 105

```

```

Db 256 -----CTGCCGGTCCAGCGCTCCACCGCTGCCACCTGCTCTGA 296
QY nThrProHISLeuValIGlyProIleArgValPheLeuAspGluPro---AspPheLY 124
Db 297 CTGGCAGCTGTGTTGGGAGACTGTGATCGATTCAACCTCAACCCATGCCCTGGA 356
QY sThrLeuGluValIleTyProAspThrHisAlaGlyIleGlyMetProLYAspCY 144
Db 357 GCGGGTGAGAGGAGAACCCAGCGCTGCTCATGGCGCGCATACACACCGCCGAGCTG 416
QY eValAlaArgHisArgValAlaIleValProTyArgAspArgGluAlaHisLeuAr 164
Db 417 CACCCAGCCAGACCGTGGCTGATCATCCCTTTAACAACCGGGAACACACCTGGCG 476
QY gLleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnIleuAspTyArgAlaIlePh 184
Db 477 CTACTGGCTCCCATCTATCTACCCCATCTTGAAGCGCGACGGCGCTGCGCTACGGGCTCA 536
QY eIleValIGluGlnValAlaAsnGlnThrPheAsnArgGlyLYsLeuMetAsnValIGlyTY 204
Db 537 TGTCAATCAACAGCATGTGTAGAGAACCTTCAACCGGCGCAAGCTGTTAAGTGGGCTT 596
QY r-----AspValAlaSerArgLeuTyProTyrGlnCysPheIlePh 218
Db 597 CCTAAGAGCGCTGAAGAGATGCCGCC-----TATGACTGCTTCATCTT 641
QY eHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThrCysProIleGluPr 238
Db 642 CAGCGATGTGAGCTGTGCTCCATGATGATGACGCAACCTATACCGGTGGGAGACCAAC 701
QY oArgHisMetSerValAlaIleAspLYsPheAsnTyArgLYsLeuProTySerAlaIlePh 258
Db 702 CGGCACCTTGGCATTTGCCATGAGACAGTTGGCTTCCGGCTTCCCTATGCTGGCTACTT 761
QY eGlyGlyIleSerAlaLeuThrLYsAspHisLeuLYsIleAsnGlyPheSerAsnAs 278
Db 762 TGGAGGTGTGTGAGCGCTGAGTAAAGCTTCAGATTTCAGAAATCAATGAGCTTCCCAATGA 821
QY pHeTrpGlyTyArgLYsGlyAspAspAspLeuAlaThrArgThrMetAlaGlyLe 298
Db 822 GATCTGGGGCTGGGGGGGAGAGATGATGACATCTTCAACCGGATCTCCCTACTGGAGAT 881
QY uLYsLeuSerArgTyProThrGlnIleAlaArgTyArgLYsMetIleLYsHisSerThrG1 318
Db 882 GAAGATCTACGCCCGCAGACATCCGAATCGCGGCTACCGCATGATCAACAGACCGGCA 941
QY uAlaThrAsnProValAsnLYsCysArgTyArgLYsIleMetGlyGlnThrLYsArgArgTr 338
Db 942 CAAAGCATAAAGAACTTAACCTCAAGGGTTTACCAAGATTCAAAAACAGAACCTGACCAT 1001
QY pThrArgAspGlyLeuSerAsnLeuLeuTyArgLYsLeuValAsnLeuGluLeuLYsProLe 358
Db 1002 GAAGGGAGCGGCTATGGGTGATGACAGCATGCTTGAAGGTGTCTCGGCAACCAT 1061
QY uTyArgThrArgAlaValAlaAspLeu 366
Db 1062 CTTCAACCAATATCAAGTGCACATT 1086

```

```

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-652A-33

Alignment Scores:
Pred. No.: 3,2e-75 Length: 1116
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
DB: 3 Gaps: 6

US-10-661-430-1 (1-383) x US-10-132-652A-33 (1-1116)
QY 7 AlAValAlaArgLeuLYsSerLeuLeuValLeuCysAlaValLeuLeuValHisAl 26
Db 24 GCTGAGGCGGCTGTGCAAGGCTGTGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 83
QY 26 aMetIleTyArgLYsIleProSerLeuTyArgLYsAsnLeuThrIleGlySerSer---Thle 45
Db 84 CATCTCTAC---TTGACGTCTACGCCAGCACCTGGACCTTCTTCAACCGCTTCAGTGC 140
QY 45 uIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAspAs 65
Db 141 CCGAGGCCCTGCCCATGCTCTCCACCCAGCTGTGACAGCAGCAGCAGCAGCAGCAGCAG 200
QY 65 pLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGlyValaAsnGlnThrSerPh 85
Db 201 CTCCCGGCCCAAGCCAGCCGCTGTAGCTCCGGGCTCCCTGAGGCCAGTGC----- 255
QY 85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnG1 105
Db 256 -----CTGCCGGTCCCAAGCTCCACCGCTCCACCGCTCCACCGCTTCCTGTA 296
QY 105 nThrProHISLeuValIGlyProIleArgValPheLeuAspGluPro---AspPheLY 124
Db 297 CTGGCAGCTGTGTTGGGAGACTGTGATCGATTCAACCCATGCCCTGGA 356
QY sThrLeuGluValIleTyProAspThrHisAlaGlyIleGlyMetProLYAspCY 144
Db 357 GCGGGTGAGAGGAGAACCCAGCGGTGCTCATGCTGAGCGCGCATACACACCGCCGAGCTG 416
QY eValAlaArgHisArgValAlaIleIleValProTyArgAspArgGluAlaHisLeuAr 164
Db 417 CACCCAGCCAGACCGTGGCGGTGATCATCCCTTTAGACACCGGGAACACCACTGCG 476
QY gLleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnIleuAspTyArgAlaIlePh 184
Db 477 CTACTGGCTCCCATCTATCTACCCCATCTTGAAGCGCGACGGCGCTGCGCTACGGGCTCA 536
QY eIleValIGluGlnValAlaAsnGlnThrPheAsnArgGlyLYsLeuMetAsnValIGlyTY 204
Db 537 TGTCAATCAACAGCATGTGTAGAGAACCTTCAACCGGCGCAAGCTGTTAAGTGGGCTT 596
QY r-----AspValAlaSerArgLeuTyProTyrGlnCysPheIlePh 218
Db 597 CCTAAGAGCGCTGAAGAGATGCCGCC-----TATGACTGCTTCATCTT 641
QY eHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThrCysProIleGluPr 238
Db 642 CAGCGATGTGAGCTGTGCTCCATGATGATGACGCAACCTATACCGGTGGGAGACCAAC 701
QY oArgHisMetSerValAlaIleAspLYsPheAsnTyArgLYsLeuProTySerAlaIlePh 258
Db 702 CGGCACCTTGGCATTTGCCATGAGACAGTTGGCTTCCGGCTTCCCTATGCTGGCTACTT 761
QY eGlyGlyIleSerAlaLeuThrLYsAspHisLeuLYsIleAsnGlyPheSerAsnAs 278
Db 762 TGGAGTGTGTGAGCGCTGAGTAAAGCTCAAGTTCTGAGAAATCAATGAGCTTCCCAATGA 821

```

```

Qy 278 pPhetPglYTrpGlyGlyValuAspAspAlaThrArgThrsSerMetAlaGlyLe 298
Db 822 GTRCTGGGGCTGGGGTGGCGAGATGATGATCATCTTCAACCGGATCTCCCTGACTGGAG 881
Qy 298 ulysValSerArgTyProThrglnlleAlaArgTyLysMetIleLysHisSerThrgl 318
Db 882 GAAGATCTACGCGCCAGACATCCGAATCCGCCCTACCGCATGATCAAGACACCGCGCA 941
Qy 318 ulatThrAsnProValAsnLysSerArgTyLysIleMetGlyGlnThrArgArgTr 338
Db 942 CAAGCATTAACGAACCTTAACCTTCAAGATTACCAAGATTCAAAACACGAAGCTGACCAT 1001
Qy 338 pThrArgAspGlyLysSerAsnLeuLysTyLysLeuValAsnLeuGlnLeuLysProle 358
Db 1002 GAAGCGGAGACGATTTGGTGGTCAAGTGGGTACAGGTCTTGAGAGTGTCTCGGACACCT 1061
Qy 358 uTyThrArgAlaValAlaAspLeu 366
Db 1062 CTTACCAATATATCACAGTGACATT 1086

RESULT 9
US-09-949-016-2195
; Sequence 2195, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2195
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2195

Alignment Scores:
Pred. No.: 3,21e-75 Length: 1119
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
Gaps: 6

US-10-661-430-1 (1-383) x US-09-949-016-2195 (1-1119)
Qy 7 AlAValAlaArgLeu-LysSerLeuLeuValLeuCyAlaValLeuLeuValHisAl 26
Db 24 GCTGAGCGGCTCTGCAAGGTGTGCTCTTCTGCTGTGTGACATTCCTCGTGGCGCT 83
Qy 26 aMetIleTyLysIleProSerLeuTyGluAsnLeuThrIleGlySerSer---Thrle 45
Db 84 CATCTCTAC---TTTGACGTCTACGCCACGACCTTGCTTTCAGCCGCTTCACTGTC 140
Qy 45 ulleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrLaseThrIserAspAs 65
Db 141 CCGAGCGCCCTGCGCCATCTTCCACCCAGCTGTAAGCGAGCGACGACGACGACCACTG 200
Qy 65 pLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerph 85
Db 201 CTCCCGGCGCAACGACCGCTCTAGTCCGGGCTCCCTGAGGTCCCACTGCC----- 255
Qy 85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCyAsnG 105

```

```

Db 256 -----CTGCCGCTCCACGCGCTCCACGCTCCACGCTCCGCTGCTCTGA 296
Qy 105 nThrProHIsleuValGlyProIleArgValPheLeuAspGluPro---AspPheLys 124
Db 297 CTGGCCACCTGGTCTGTGGGAGAGTGTGATCGAGTTCACTTACCTACCTACCGATCCCTGGA 356
Qy 124 sThrLeuGluLysIleTyProAspThrHisAlaGlyGlnHisGlyMetProLysAspCy 144
Db 357 GCGGGTGCAGAGGAGGAACCAAGCGGTGCTCATGGCGCGCGCATACACACCGCGCGACTG 416
Qy 144 sValAlaArgHisArgValAlaIleLeuValProTyThrArgAspArgGluAlaHisLeuAr 164
Db 417 CACCCCGACCCAGACCGGTGGGTATCTCCCTTTTAAACACCGGGAACACCACTGGG 476
Qy 164 gIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyThrAlaIleph 184
Db 477 CTACTGGCTTCACATATCTACACCCCATCTTGAGCGCGGACGCGGCTCGCTACGGCGCTTA 536
Qy 184 eIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTy 204
Db 537 TGTCAATCAACGACATGATGAGGACACTTCAACGCGGCAAGCTCTTAAGTGGGCTT 596
Qy 204 r-----AspValAlaSerArgLeuTyProTyPglncCyPheIleph 218
Db 597 CTTAGAGCGGCTGAAGAGATGCCGCC-----TATGACTGCTTCACTT 641
Qy 218 eHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThrCyAspProIleGlnPr 238
Db 642 CAGCATGTGACCTGTCTCCCATGTGATACCGCAACTTATACCGTGGCGGACCAACC 701
Qy 238 oArgHisMetSerValAlaIleAspLysPheAsnTyLysLeuProTySerAlaIleph 258
Db 702 CGGCACCTTGGCATTTGCATGTGACGACAAATTGGCTTCCGGCTTCCCTATGCTGGTACTT 761
Qy 258 eGlyGlyLysSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAs 278
Db 762 TGGAGGTGTGTCCAGGCTGATGAAGCTCATGTTCTGAANAATCAATGCGCTTCCCATATGA 821
Qy 278 pPhetPglYTrpGlyGlyValuAspAspAlaThrArgThrsSerMetAlaGlyLe 298
Db 822 GTRCTGGGGCTGGGGTGGCGAGATGATGATCATCTTCAACCGGATCTCCCTGACTGGAG 881
Qy 298 ulysValSerArgTyProThrglnlleAlaArgTyLysMetIleLysHisSerThrgl 318
Db 882 GAAGATCTACGCGCCAGACATCCGAATCCGCCCTACCGCATGATCAAGACACCGCGCA 941
Qy 318 ulatThrAsnProValAsnLysSerArgTyLysIleMetGlyGlnThrArgArgTr 338
Db 942 CAAGCATTAACGAACCTTAACCTTCAAGATTACCAAGATTCAAAACACGAAGCTGACCAT 1001
Qy 338 pThrArgAspGlyLysSerAsnLeuLysTyLysLeuValAsnLeuGlnLeuLysProle 358
Db 1002 GAAGCGGAGACGATTTGGTGGTCAAGTGGGTACAGGTCTTGAGAGTGTCTCGGACACCT 1061
Qy 358 uTyThrArgAlaValAlaAspLeu 366
Db 1062 CTTACCAATATATCACAGTGACATT 1086

RESULT 10
US-09-949-016-2196
; Sequence 2196, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08

```



```

QY 105 nThrProtrOnHLeuValGlyProIleArgValPheLeuAspGluPro---AspPheLeu 124
Db 466 CTGGCCACCTGGCTTGTGGGAGAGACTGGTGAAGTTGACCTGACCCATGGCCCTGGGA 525
QY 124 sThrLeuGluValIleTyProAspThrThiValAGlyGlyHisGlyMetProLysAspCy 144
Db 526 GCGGGGTGCAGAGAGAGAACCCAGAGCGTGCATATGGGGCGCGGATACACACCGCCGACTG 585
QY 144 sValAlaArgHisArgValAlaIleIleValProTyArgAspArgGluAlaHisLeuArg 164
Db 586 CACCCCAAGCCCAAGCGGTGGGTGCATCATCTCCCTTGAACACCGGACACCACTGGCG 645
QY 164 gileMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyraIleLeu 184
Db 646 CTACTGGCTCCATATTTACACCCCACTTGGAGCGGCGACCGGCTGGCGCTACGGCGCTTA 705
QY 184 eIleValAGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTy 204
Db 706 TGTCAATCAACCAAGCATGTGAGAGACACTTCAACCGGGCCCAAGCTGCTTAACGTGGGCTT 765
QY 204 r-----AspValAlaSerArgLeuTyProTrpGlnCysPheIleLeu 218
Db 766 CTTAGAGCGCGCTGAAGAGAGATCCGCC-----TATGACTGCTTCACTT 810
QY 218 eHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThrCysProIleGlnPr 238
Db 811 CACGAGATGTGACTGTGTCCCATGTGATGACCGGAACCTTATCCGCGCGGCGACCAACC 870
QY 238 oArgHisMetSerValAlaIleAspLysPheAsnTyThrLysLeuProTySerAlaIlePh 258
Db 871 CCCCACCTTTCGCAATGCCATGACAGAGTTGGCTTCGGGCTTCCTATCTCGGCTACTT 930
QY 258 eGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAs 278
Db 931 TGAAGGTGTCTCAGGCGCTGAGTACAGCTCTTCTAGATCATAGCTTCCCAATGA 990
QY 278 pPheTrpGlyTyTrpGlyGlyGluAspAspLeuAlaThrArgThrSerMetAlaGlyLe 298
Db 991 GTACTGGGGGTGGGTGGCGGAGTATGATGATCTTCAACCGGATCTCCCTGACTGGGAT 1050
QY 298 uLysValSerArgTyProThrGlnIleAlaArgTyLysMetIleLysHisSerThrG1 318
Db 1051 GAAGATCTCAGCCGACAGACATCCCAATCGCGCTACCGCATGATCAAGACAGACCGCGA 1110
QY 318 uAlaThrAsnProValAsnLysCysArgTyLysIleMetGlyGlnThrLysArgTr 338
Db 1111 CAAGCATTAAGCAACTAACCTCAGAGGTTTACCAAGATTCAAAACAGAAAGCTGACCAT 1170
QY 338 pThrArgAspGlyLeuSerAsnLeuLysTyLysLeuValAsnLeuGluLeuLysProLe 358
Db 1171 GAAGCGGAGACGCAATTGGGTCAAGTGGCGTACAGGTCTTGAGAGGTCTCGGCAACCACT 1230
QY 358 uTyThrArgAlaValAlaAspLeu 366
Db 1231 CTTCAACCAATATCAAGTGGACATT 1255

```

```

; LENGTH: 2027
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-652A-1
Alignment Scores:
Pred. No.: 8,74e-75 Length: 2027
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32,408 Indels: 23
DB: Gaps: 6
US-10-661-430-1 (1-383) x US-10-132-652A-1 (1-2027)
QY 7 AlaValAlaArgLeu-LysSerLeuLeuValLeuCysValAlaValLeuLeuValHisAl 26
Db 193 GCTGAGCGCGCTGTGAGAGCTGTGCTCTTCTGTGCTGTGCACTTCTCTGGCGCT 252
QY 26 aMetIleTyLysIleProSerLeuTyArgLysAsnLeuThrIleGlySerSer---Thre 45
Db 253 CATCTCTAC---TTGAGGTCTACGCCAGACAGCACTGGCGCTTCTTCAAGCGCTTACGTC 309
QY 45 uIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAspAs 65
Db 310 CCGAGGCGCTGCCCATGCGCTTACCGCCAGCTGTAAGACAGCAGCAGCAGCAGCAACTG 369
QY 65 pLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGlyValAlaAsnGlnThSerPh 85
Db 370 CTCCCGGCCCAACGCCACCGCTTACGTCCGGGCTCCCTGAGGTCCCAAGTCC----- 424
QY 85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnG1 105
Db 425 -----CTGCCGCGTCCCAACCGCTCCCAAGCTCCACCGCTCCACCGTCTCGA 465
QY 105 nThrProtrOnHLeuValGlyProIleArgValPheLeuAspGluPro---AspPheLeu 124
Db 466 CTGGCCACCTGGCTTGTGGGAGAGACTGGTGAAGTTGACCTGACCCATGGCCCTGGGA 525
QY 124 sThrLeuGluValIleTyProAspThrThiValAGlyGlyHisGlyMetProLysAspCy 144
Db 526 GCGGGGTGCAGAGAGAGAACCCAGAGCGTGCATATGGGGCGCGGATACACACCGCCGACTG 585
QY 144 sValAlaArgHisArgValAlaIleIleValProTyArgAspArgGluAlaHisLeuArg 164
Db 586 CACCCCAAGCCCAAGCGGTGGGTGCATCATCTCCCTTGAACACCGGGAACACCACTGGG 645
QY 164 gileMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyraIleLeu 184
Db 646 CTACTGGCTCCATATTTACACCCCACTTGGAGCGGCGACCGGCTGGCGCTACGGCGCTTA 705
QY 184 eIleValAGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTy 204
Db 706 TGTCAATCAACCAAGCATGTGAGAGACACTTCAACCGGGCCCAAGCTGCTTAACGTGGGCTT 765
QY 204 r-----AspValAlaSerArgLeuTyProTrpGlnCysPheIleLeu 218
Db 766 CTTAGAGCGCGCTGAAGAGATCCGCC-----TATGACTGCTTCACTT 810
QY 218 eHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThrCysProIleGlnPr 238
Db 811 CACGAGATGTGACTGTGTCCCATGTGATGACCGGAACCTTATCCGCGCGGCGACCAACC 870
QY 238 oArgHisMetSerValAlaIleAspLysPheAsnTyThrLysLeuProTySerAlaIlePh 258
Db 871 CCCCACCTTTCGCAATGCCATGACAGAGTTGGCTTCGGGCTTCCTATGCTGACTTACTT 930
QY 258 eGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAs 278
Db 931 TGAAGGTGTCTCAGGCGCTGAGTACAGCTTCTTGAAGATCAAGAGGTCTCCCAATGA 990
QY 278 pPheTrpGlyTyTrpGlyGlyGluAspAspLeuAlaThrArgThrSerMetAlaGlyLe 298
Db 991 GTACTGGGGGTGGGTGGCGGAGTATGATGATCTTCAACCGGATCTCCCTGACTGGGAT 1050

```

```
Db 991 GTAGTGGGCTGGGGTGGCGAGATGATGACATCTTCAACCGGATCTCCCTGACTGGGAT 1050
Qy uUyValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleValHisSerThrG1 318
Db 1051 GAAGATCTACCGCCAGACATCCGAATCGCGCTACCGCATGATCAAGACAGACCGGCA 1110
Qy 318 uAlaThrAsnProValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTr 338
Db 1111 CAAGATAAACGAACCTTAACCTTCAGAGCTTTCACAAAGATTCAAAACAGAACCTGACCT 1170
Qy 338 pThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProle 358
Db 1171 GAAGCGGACCGGCTGCTGAGTGGGTACAGGCTTGAGAGTGCTCGGCAACCACT 1230
Qy 358 uTyrThrArgAlaValAlaAspLeu 366
Db 1231 CTTCAACCAATATCAGATGACATT 1255

RESULT 13
US-09-118-464-34
; Sequence 34, Application US/09118464A
; Patent No. 6558934
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Bennett, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
; TITLE OF INVENTION: Beta-1,4-galactosyltransferase, BetaGal-T2
; FILE REFERENCE: 4305/08521
; CURRENT APPLICATION NUMBER: US/09/118,464A
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-118-464-34

Alignment Scores:
Pred. No.: 1,01e-73 Length: 1023
Score: 646.00 Matches: 120
Percent Similarity: 61.65% Conservative: 52
Best Local Similarity: 43.01% Mismatches: 95
Query Match: 31.81% Indels: 12
Gaps: 3

US-10-661-430-1 (1-383) x US-09-118-464-34 (1-1023)
Qy 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValAlaIleProIle 114
Db 172 CCCAGGCTCCCAACGCTGACCCCTGCTGACTGCGCCACCTGGTCTTGTGGGCAAGACTG 231
Qy 115 ArgValPheLeuAspGluPro--AspPheLysThrLeuGluLysIleTyrProAspThr 133
Db 232 CTGATCGAGTTCACCTCAACCTCAAGCCCTGAGCGGGTCAAGAGGAGAACCAAGCGCTG 291
Qy 134 HisAlaGlyGlnHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 292 CTCATGGGGGCGGATACACACCGCCGACTGACCCCAAGCGGATGGCGGTCATC 351
Qy 154 ValProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
Db 352 ATCCCTTTAGACACCGGAGAACACCACTGCGCTGCTGCTCACTATTAACACCCCACT 411
Qy 174 LeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
Db 412 TTGAGGCGGACAGCGGCTGCGCTACCGGCTGTATGTCATCAACACAGATGTAGAGACACC 471
Qy 194 PheAsnArgGlyLysLeuMetAsnValGlyTyr-----AspValAla 207
Db 472 TTCACACCGGCGCAAGGCTGTAAAGTGGCTTCCCTTAAGAGCGCTGAAGAGATGGCGCC 531
Qy 208 SerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAspLeuLeuProGluAsp 227
```

```
Db 532 -----TATGACTGCTTCATCTTCAAGCGATGAGACTGGTCCCATGAT 576
Qy 228 AspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSerValAlaIleLysLys 247
Db 577 GACCGCAACCTTAACCGCTGGCGGACACCGCCGCACTTGGCCATTGCCATGGACAG 636
Qy 248 PheAsnTyrLysLeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAsp 267
Db 637 TTGGCTTCCGCTTCCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 696
Qy 268 HisLeuLysLysIleAsnGlyPheSerAsnAspPheTrpGlyTyrGlyValLysAsp 287
Db 697 CAGTTTCAGATTAATAGGCTTCCCAATGATGACTGGGGCTGGGGGCGGAGATGAT 756
Qy 288 AspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIle 307
Db 757 GACATCTTCAACCGGATCTCCCTGACTGGATGAAGATCTCAACGCCAGACATCCGAATC 816
Qy 308 AlaArgTyrLysMetIleLysHisSerThrGlnAlaThrAsnProValAsnLysCysArg 327
Db 817 GCGCGCTACCGATGATCAAGACAGCCGCAAGCAATTAACGAACCTTAACCTTCAAG 876
Qy 328 TyrLysIleMetGlyGlnThrLysArgArgTrpThrArgAspGlyLeuSerAsnLeuLys 347
Db 877 TTTACCAAGATTCAAAACAGAGCTGACATGAAGCGGAGCGGATGGGTGAGTGGCGG 936
Qy 348 TyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
Db 937 TACCAGTCTTGGAGTGTCTCGGCACACCTTCAACCAATATCAAGTGAACATT 993

RESULT 14
US-10-132-652A-34
; Sequence 34, Application US/10132652A
; Patent No. 6916649
; GENERAL INFORMATION:
; APPLICANT: CLAUSEN, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; TITLE OF INVENTION: 4-GALACTOSYLTRANSFERASE, B4 GAL-T2
; FILE REFERENCE: 04305/1008521-US1
; CURRENT APPLICATION NUMBER: US/10/132,652A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 09/118,464
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 1023
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-132-652A-34

Alignment Scores:
Pred. No.: 1,01e-73 Length: 1023
Score: 646.00 Matches: 120
Percent Similarity: 61.65% Conservative: 52
Best Local Similarity: 43.01% Mismatches: 95
Query Match: 31.81% Indels: 12
Gaps: 3

US-10-661-430-1 (1-383) x US-10-132-652A-34 (1-1023)
Qy 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValAlaIleProIle 114
Db 172 CCCAGGCTCCCAACGCTGACCCCTGCTGACTGCGCCACCTGGTCTTGTGGGCAAGACTG 231
Qy 115 ArgValPheLeuAspGluPro--AspPheLysThrLeuGluLysIleTyrProAspThr 133
Db 232 CTGATCGAGTTCACCTCAACCTCAAGCCCTGAGCGGGTCAAGAGGAGAACCAAGCGCTG 291
Qy 134 HisAlaGlyGlnHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 292 CTCATGGGGGCGGATACACACCGCCGACTGACCCCAAGCGGATGGCGGTCATC 351
Qy 352 ATCCCTTTAGACACCGGAGAACACCACTGCGCTGCTGCTCACTATTAACACCCCACT 411
Db 412 TTGAGGCGGACAGCGGCTGCGCTACCGGCTGTATGTCATCAACACAGATGTAGAGACACC 471
Qy 472 TTCACACCGGCGCAAGGCTGTAAAGTGGCTTCCCTTAAGAGCGCTGAAGAGATGGCGCC 531
Qy 531 134 HisAlaGlyGlnHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
```

Db 292 CTCATGGGCGCGCATATACACCCGCCGATCGACCCAGCCGATGCGCGTCAATC 351
Qy 154 ValProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
Db 352 ATCCCTTTAGACACCGGAAACACACTGCTACTGCTGCTCAATCTACACCCCATC 411
Qy 174 LeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
Db 412 TTGAGCGCGAGCGGCTCGGCTCAAGGCTCATGTATCAACAGATGTGTAGACAC 471
Qy 194 PheAsnArgGlyLysLeuMetAsnValGlyTyr-----AspValAla 207
Db 472 TTCACCGCGGCGCACTGCTTAACTGGGCTCTAGAGCGCGCTGAAGAGATGCCGCC 531
Qy 208 SerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAspLeuLeuProGluAsp 227
Db 532 -----TATGACTGCTTATCTTACGCGATGTGAGCTGGTCCCATGTGAT 576
Qy 228 AspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLys 247
Db 577 GACCGCAACTATACCGCTGCGGAGCAACCCGCCACTTGGCATTGCCATGGCAGACAG 636
Qy 248 PheAsnTyrLysLeuProTyrSerAlaIlePheGlyGlyLysSerAlaLeuThrLysAsp 267
Db 637 TTGGGCTTCCGCGCTCCCTATGCTGCTGCTACTTGGAGGTGTGTACAGGCTGAGTAAAGGCT 696
Qy 268 HisLeuLysLysIleAsnGlyPheSerAsnAspPheThrGlyTyrThrGlyGlyIleAspAsp 287
Db 697 CAGTTCTGGAATCAATGAGCTTCCCATGAGACTGGGGCTGGGGTGGCGAGAGAT 756
Qy 288 AspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIle 307
Db 757 GACATCTTCAACCGGATCTCCCTGACCTGGGATGAAGATCTCAGCCGACATCCGATC 816
Qy 308 AlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArg 327
Db 817 GGCCCGCTACCGCATGATCAAGACGACCGGCAAGCATACGAACTTAACCTCAGAGG 876
Qy 328 TyrLysIleMetGlyGlnThrLysArgTyrThrArgAspGlyLysSerAsnLeuLys 347
Db 877 TTTCACCAAGATTCAAAACAGAGAGCTACCATGAAGCGGAGCGGATGGGTCACTCGG 936
Qy 348 TyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAlaValIleAspLeu 366
Db 937 TACAGGCTTGGAGGTGTCTCGGCAACCACTTTCACCAATATCAACAGTGCACATT 993

RESULT 15
US-09-991-181-235
Sequence 235, Application US/09991181
Patent No. 6913919
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary B.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gunney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
Acids Encoding the Same
FILE REFERENCE: P2730PIC53
CURRENT APPLICATION NUMBER: US/09/991,181
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088861
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/088876
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089105
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/089440
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089512
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089514
 PRIOR FILING DATE: 1998-06-16
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089538
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089598
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089600
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089653
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089801
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089908
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090676
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090678
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090690

PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090694
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090695
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090696
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090862
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091478
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091544
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091626
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091633
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091978
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/092182
 PRIOR FILING DATE: 1998-07-09

Alignment Scores:
 Pred. No.: 8,65e-73 Length: 1964
 Score: 642.50 Matches: 143
 Percent Similarity: 53.37% Conservative: 55
 Best Local Similarity: 38.54% Mismatches: 124
 Query Match: 31.63% Indels: 49
 Gaps: 8

US-10-661-430-1 (1-383) x US-09-991-181-235 (1-1964)

QY 5 HisLeuAlaValAla--ArgLeuLysSerLeuLeuValLeuValLeuLeu--- 22
 DB 140 CACCTTCTCAAAATTCGATTCGCTGTTGACTTGTGCTGACCTGTGCTGACAGTGTGGG 199
 QY 23 -----LeuValHisAlaMetIleTyrLysIleProSerLeuTyrGlu 36
 DB 200 TGGGCCACCACTACTCTGCTGGTGGTCC--ATTCAAGAGATTCTTAAGCAAAAGAG 256
 QY 37 AsnLeuThrIleGlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeu 56
 DB 257 TTCATGGCTAATTCCTCAATTAAGCCCTCATTTTG-----GGGAAGGAAAAACTCTG 307
 QY 57 GlyAsnThrAlaSerThrSerAspAspLeuLeuAspThrTTPAsnSerThrPheSerPro 76
 DB 308 ACTAATGAAGCATCCACGAAGAGTAGAAGCTTGACAAAC----- 346
 QY 77 IleSerGluValAsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAsp 96
 DB 346 ----- 346
 QY 97 AsnGlnThrLeuGlnPheCysAsnGlnThrProPheIleValGlyProIleArgVal 116
 DB 347 -----TGTCTTCTGTGTCTCTTAAGCTTAAGAGGCAAGCAAGCTC 388
 QY 117 PheLeuAspGluProAspPheLysThrLeuGluLysIleTyrProAspThrHis----- 134
 DB 389 ATTTTC--AAACCAATCTC--ACTTGAAGAGGTACAGGCAAGAAATCCCAAGTG 442
 QY 135 AlaGluGlyHisGlyMetProLysAspCysValAlaGlnHisArgValAlaIleIleVal 154
 DB 443 TCACAGAGCCGGTATCGCCCTCAGGAATGTAAAGCTTACAGAGAGGTGCGCATCTCGTT 502
 QY 155 ProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeu 174

Db	503	CCCCACCGGAAACAGAGAGAAACACCTGATCTGACCTGCGAAACATCTGCATCCCTTCTG	562
Qy	175	AlAlsgInglInleuAspTyrAlaAlIephellValGluInValAlaAsnGlnThrPhe	194
Db	563	CAGAGCGAGCGAGCTGATTTATGCGATCTACCTCATCCACCGAGGTGAAGTAAAGATT	622
Qy	195	AspArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTrogIn	214
Db	623	AATGAGCGCAAACTCTTGATGATGTGGCGCTATCTAGAAGCCCTCCAGGAAGAAATTGGAC	682
Qy	215	CysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCys	234
Db	683	TGCTTTTATTTCCAGATGTGAGACCTGGTACCCCGAATGACTTTAACTTTACAAAGTGT	742
Qy	235	ProIleGlnProAlaGHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyr	254
Db	743	GAGAGCATCCCAAGACATCTGTGTGTGGCAGGAACAGCACTGGGTACAGGTACGTTAC	802
Qy	255	SerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLysAlaAsnGly	274
Db	803	AGTGGATATTTTGGGGGGTGTACTCTCCCTAAGCAGAGAGCAGTTTTCAGGTGAATGGA	862
Qy	275	PheSerAsnAspPheMetArgLysTyrGlyGlyGlyLysAspAspLeuAlaThrArgThrSer	294
Db	863	TTCCTTAACACTACTGTGGAGATGGGAGAGCGAAGCGATGACCTCAGACTCAGGGTTGAG	922
Qy	295	MetAlaGlyLeuLysValIleSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLys	314
Db	923	CTCCAAAGATGAATAATTTCCCGGCGCCCTGCCTGAGAGGGGTAAATATACATAGTCTTC	982
Qy	315	HisSerThrGluAlaThrAspProValAsnLysCysAspTyrLysIleMetGlyGlnThr	334
Db	983	CACACTAGACACAAAGCAATGAGGTGAACCGACAGACGATGAGAGCTTTACACCAAGTGTG	1042
Qy	335	LysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGln	354
Db	1043	TCACGAGCTCGGAGAAACAGATGGGTGAGTGTCTTTAAATTGATCTGTGGAA	1102
Qy	355	LeuLysProLeuTyrThrArgAlaValValAsp	365
Db	1103	CACAACTCCTTATATATCAACATCAACATCGAT	1135

Search completed: January 9, 2006, 04:25:47
Job time : 260 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: January 9, 2006, 03:00:11 ; Search time 988 Seconds
(without alignments)
3205.639 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAAVARKSLVLCV.....VDLKEKCRRLRDPPTCF 383

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 9793542 segs, 413468905 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seg length: 0
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=nlh
-Q=/cgn2_1/USPTO_spool/US10661430/runat_06012006_123138_6363/app_query.fasta_1.519
-DB=Published_Applications_NA_Main -QFMT=fastap -SUFFIX=p2n.rnpbm
-MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1
-MATRIX=10sum62 -TRANS=human40.cdi -LIST=45 -DOCLAM=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10661430 @CCN_1_1.1026 @runat_06012006_123138_6363 -NCRU=6 -ICPU=3
-NO_MAP -LARGEDEBERT -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCAPEXT=7 -YGAPOP=7 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA_Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	1152	7	US-10-661-430-2
2	718	35.4	2002	3	US-09-925-301-230
3	709	34.9	1839	10	US-11-097-143-18554
4	684	33.7	1853	3	US-09-745-763-84
5	675	33.2	4093	10	US-11-097-143-18553
6	667.5	32.9	3953	5	US-10-087-192-2048
7	666	32.8	1158	8	US-10-466-941-5

8	666	32.8	1158	9	US-10-870-635-5	Sequence 5, Appl1
9	666	32.7	4162	5	US-10-084-817-168	Sequence 168, App
10	663.5	32.7	1116	8	US-10-761-435A-19	Sequence 19, Appl
11	661	32.5	1119	7	US-10-287-226-81	Sequence 81, Appl
12	660.5	32.5	1191	9	US-10-713-970-20	Sequence 20, Appl
13	658	32.4	1116	5	US-10-132-652-33	Sequence 33, Appl
14	658	32.4	1116	10	US-11-105-796-33	Sequence 33, Appl
15	658	32.4	1890	9	US-10-956-157-1114	Sequence 114, App
16	658	32.4	2027	5	US-10-132-652-1	Sequence 1, Appl1
17	658	32.4	2027	10	US-11-105-796-1	Sequence 1, Appl1
18	657.5	32.4	3194	8	US-10-723-860-7737	Sequence 7737, Appl
19	655.5	32.3	6406	9	US-10-713-970-7	Sequence 7, Appl1
20	655	32.3	4080	3	US-09-960-706-487	Sequence 487, App
21	655	32.3	4080	3	US-09-873-319-297	Sequence 297, App
22	646	31.8	1023	5	US-10-132-652-34	Sequence 34, Appl
23	646	31.8	1023	10	US-11-105-796-34	Sequence 34, Appl
24	643.5	31.7	1888	7	US-10-287-226-77	Sequence 77, Appl
25	642.5	31.6	1964	3	US-09-989-722-235	Sequence 235, App
26	642.5	31.6	1964	3	US-09-989-723-235	Sequence 235, App
27	642.5	31.6	1964	3	US-09-989-723-235	Sequence 235, App
28	642.5	31.6	1964	3	US-09-989-727-235	Sequence 235, App
29	642.5	31.6	1964	3	US-09-989-727-235	Sequence 235, App
30	642.5	31.6	1964	3	US-09-989-732-235	Sequence 235, App
31	642.5	31.6	1964	3	US-09-991-073-235	Sequence 235, App
32	642.5	31.6	1964	3	US-09-990-442-235	Sequence 235, App
33	642.5	31.6	1964	3	US-09-991-163-235	Sequence 235, App
34	642.5	31.6	1964	3	US-09-990-604-235	Sequence 235, App
35	642.5	31.6	1964	3	US-09-990-456-235	Sequence 235, App
36	642.5	31.6	1964	3	US-09-989-721-235	Sequence 235, App
37	642.5	31.6	1964	3	US-09-992-558-235	Sequence 235, App
38	642.5	31.6	1964	3	US-09-989-293A-235	Sequence 235, App
39	642.5	31.6	1964	3	US-09-989-735-235	Sequence 235, App
40	642.5	31.6	1964	3	US-09-990-444-235	Sequence 235, App
41	642.5	31.6	1964	3	US-09-991-181-235	Sequence 235, App
42	642.5	31.6	1964	3	US-09-989-730-235	Sequence 235, App
43	642.5	31.6	1964	3	US-09-990-436-235	Sequence 235, App
44	642.5	31.6	1964	3	US-09-993-687-235	Sequence 235, App
45	642.5	31.6	1964	3	US-09-989-734-235	Sequence 235, App

ALIGNMENTS

RESULT 1
US-10-661-430-2
; Sequence 2, Application US/10661430
; Publication No. US20040086995A1
; GENERAL INFORMATION:
; APPLICANT: Cummings, Richard D.
; TITLE OF INVENTION: Kawai, Ziad S.
; TITLE OF INVENTION: NUCLEIC ACIDS AND METHODS OF USE THEREOF
; FILE REFERENCE: 7148.001
; CURRENT APPLICATION NUMBER: US/10/661,430
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 60/411,242
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-661-430-2

Alignment Scores:
Pred. No.: 4.58e-257 Length: 1152
Score: 2031.00 Matches: 383
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0
US-10-661-430-1 (1-383) x US-10-661-430-2 (1-1152)

```
QY 1 MetAlaPheArgHisLeuAlaValAlaArgLeuLeuSerLeuLeuValLeuCysAlaVal 20
Db 1 ATGGCTTTTTCGATTTGGAGCGATCGCAGACTCAAGTCGTTGCTGACTCTTGAGCCGTT 60
QY 21 LeuLeuLeuValHisAlaMetCileTYrLYsIleProSerLeuTYrGluAsnLeuThrIle 40
Db 61 CTTCTATTAGTTCATGCAATGATTATTAAAGATTCATCGCTTACAGAAACCTTACTATAC 120
QY 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla 60
Db 121 GGCTCCTCGACCTTATTGCGCAGCGTCGACGCAATGAGCAATGCTCGGGAATACGGCT 180
QY 61 SerThrSerAspAspLeuAspThrTrpAsnSerThrPheSerProIleSerGluVal 80
Db 181 TCCACTTCGGATGATCTACTTGTGATCGTGAAATTCACGTTTCCACCGAATTTCTGAAGTT 240
QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
Db 241 AATCAGACTAGTTTATTAGAGGACATTCGTCCAATCGTGTCCCGCAACACGACTCTT 300
QY 101 GlnPheCysAsnGlnThrProTrophLeuValGlyProIleArgValPheLeuAspGlu 120
Db 301 CAATTCTGTAATCAGACACCTCCCACTCGTGGACCAATCCGTAATTCCTCGATGAG 360
QY 121 ProAspPheLeuThrLeuGluLYsIleTYrProAspThrHisAlaGlyGlyHisGlyMet 140
Db 361 CCGGACTTCGAAACTCGAGAAATTTATCCGACACGACCGCGGTGACATGGAATG 420
QY 141 ProLYsAspCysValAlaArgHisArgValAlaIleIleValProTYrArgAspArgGlu 160
Db 421 CTTAAGGATTTGTGTTGCAAGGATCGTGTGCTATTATTGCGCCCTATAGATCGTGAA 480
QY 161 AlaHisLeuArgGluMetLeuHisAsnLeuHisSerLeuLeuAlaLYsGlnIleAsp 180
Db 481 GCACATTTGAGAAATGATGCTCCACAAATTTGCATCGTCCGCAAAACAACATTTGAC 540
QY 181 TYrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLYsLeuMet 200
Db 541 TATGCAATTTTCACTATGAGCAAGTGGCAATGACGCTTAAATCCGGGAAACTAATG 600
QY 201 AsnValGlyTYrAspValAlaSerArgLeuTYrProTrpGlnCysPheIlePheHisAsp 220
Db 601 AAGCTTGATACGACGATGATGACGCTCACTCAACCATGGCAGTCTTCATCTTCATGAT 660
QY 221 ValAspLeuLeuProGluAspAspArgAsnLeuTYrThrCysProIleGlnProArgHis 240
Db 661 GTCGATTTTACTGCCCGAAGATGACGTAACCTGTACAGTGTCCAAATTCACACGTCAT 720
QY 241 MetSerValAlaIleAspLYsPheAsnTYrLYsLeuProTYrSerAlaIlePheGlyGly 260
Db 721 ATGAGTGTAGCGATCGATTAATTCATTAATACTTCATATTCGGCGATTTTCGGGGGA 780
QY 261 ILeSerAlaLeuThrLYsAspHisLeuLYsIleAsnGlyPheSerAsnAspPheTrp 280
Db 781 ATCAGTGACATAACAAAAGATCACCTGAAAGAAAATCAATGATTTTGCATTTTGG 840
QY 281 GLYrTrpGlyGlyAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLYsVal 300
Db 841 GGTGGGGGGGAGAGACGATTTGGCGAGCAAGAACATGATGATGGCTGACATAAGTT 900
QY 301 SerArgTYrProTrpGlnIleAlaArgTYrLYsMetIleLYsHisSerThrGluAlaThr 320
Db 901 TCAAGATATCCGACCAAAATTCGACGATTAATAATGATTAAAGCTCGACGGAAGCGG 960
QY 321 AsnProValaLeuLYsCysArgTYrLYsIleMetGlyGlnThrLYsArgArgTYrThrArg 340
Db 961 AATCCAGTTAATTAATGCCGCTACAAATTAATGGGCCAAACGAAAGCCGCGATGACACGT 1020
QY 341 AspGlyLeuSerAsnLeuLeuTYrLYsLeuValaLeuLeuGluLeuLYsProLeuTYrThr 360
Db 1021 GACGGCCTTAAGCAATCTGAAGTAAAGCTCGTAATCTGGAATGAAACCTCTCTACACT 1080
```

```
QY 361 ArgAlaValaIleAspLeuLeuGlyLYsAspCysArgArgGluLeuArgAspPhePro 380
Db 1081 CGAGCGCTGCTGATTTGCTGCAAAAAAGACTGCCCGCGGAGACTCGAAGGACCTTCCA 1140
QY 381 ThrCysPhe 383
Db 1141 ACGTGTATT 1149

RESULT 2
US-09-925-301-230
/ Sequence 230, Application US/09925301
/ Patent No. US20020052308A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA106
/ CURRENT APPLICATION NUMBER: US/09/925,301
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05882
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ NUMBER OF SEQ ID NOS: 1694
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 230
/ LENGTH: 2002
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-301-230

Alignment Scores:
Pred. No.: 4,14e-83 Length: 2002
Score: 718.00 Matches: 134
Percent Similarity: 64.60% Conservative: 43
Best Local Similarity: 48.91% Mismatches: 95
Query Match: 35.35% Indels: 2
DB: Gaps: 2

US-10-661-430-1 (1-383) x US-09-925-301-230 (1-2002)
QY 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIle 114
Db 470 CCAATCTCTCAAGGTGTGCCCTTACTGTCCAGAAACGATTCCTCTTAAATGGGTCTGTG 529
QY 115 ArgVal---PheLeuAspGlnProAspPheLYsThrLeuGluLYsIleTYrProAspThr 133
Db 530 TCGGTCTCTTTAGCCCAAGTGCATCATCGGACGAGATGTGAGCGGAAATCCCGGTA 589
QY 134 HisAlaGlyGlyHisGlyMetProLYsAspCysValAlaArgHisArgValAlaIleIle 153
Db 590 GAACCAAGGGGGCGGTACCGCCTGCAAGTTGTGAGCCCCCGCTCCCAACAGCCATCATTT 649
QY 154 ValProTYrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
Db 650 GTGCTTCATGTGTCCCGGAGACACACCTGCGCTGTGTGTGATCAACACTGCACCCCTTC 709
QY 174 LeuAlaLYsGlnGlnLeuAspTYrAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
Db 710 TTTCAGGCGCAGACAGCTTGCTTATGCACTATGTATCATCCACAGCGTGGAAATGGAACA 769
QY 194 PheAsnArgGlyLYsLeuMetAsnValGlyTYrAspValAlaSerArgLeuTYrProTrp 213
Db 770 TTTTAACAGGGCAAAAGCTTTGAACGTTGGGGTCCGAGAGGCCCTGCGTGAAGAAGTGG 829
QY 214 GlnCysPheIlePheHisAspValaIleAspLeuProGluAspAspArgAsnLeuTYrThr 233
Db 830 GACTGCTGTCTTTCGACGATGTGACCTTTCCAGAAATATACCAATCTGTATGTG 889
QY 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLYsPheAsnTYrLYsLeu 252
Db 890 TGTGACCCCGGGGAGACCCCGGCATGTGGCTTGATGAACAAGTTGGATACAGCTC 949
QY 253 ProTYrSerAlaIlePheGlyGlyIleSerAlaLeuThrLYsAspHisLeuLYsIle 272
```



```
Db 950 CCGTACCCCACTACTGCGAGAGCTCAGCAGCTTACTCTCGACAGAGTAAAGATG 1009
Qy 273 AenglypHeserAspAphetTTPGlyTTPGlyGluAspAspAphetAthrArg 292
Db 1010 AATGGCTTCCCAAGAAATACCTGGGGCTGGGGTGGAGATGAGACACTTCTACAGG 1069
Qy 293 ThrsMetAlaGlyLeuLeuValSerArgTyrProthGlnIleAlaArgTyrIleMet 312
Db 1070 GTGGCCCTGGCTGGAGATGATCTCTGGCCCCCAGACTCTGTGAGACACTTAAGATG 1129
Qy 313 IleYHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrIleMetGly 332
Db 1130 GTGAAGCAGCAGAGATAGAGGCAATGAGAAATCCCAAGATTGACCTCGTGC 1189
Qy 333 GlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrIleValAla 352
Db 1190 CGTACCCCAATTCCTGAGCAGCAGATGGGATGAACTCACTGACATACCACTTGGCT 1249
Qy 353 LeuGluLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
Db 1250 CGAGAGCTGGGGCTCTTTATATCAACATCAGACAGACATT 1291

RESULT 3
US-11-097-143-18554
; Sequence 18554, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18554
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-18554

Alignment Scores:
Pred. No.: 5 56e-82 Length: 1839
Score: 709.00 Matches: 150
Percent Similarity: 58.03% Conservative: 56
Best Local Similarity: 42.25% Mismatches: 129
Query Match: 34.91% Indels: 20
DB: 10 Gaps: 7

US-10-661-430-1 (1-383) x US-11-097-143-18554 (1-1839)
Qy 28 IleTyrLysIleProSerLeuTyrGluAsnLeuThr-----IleGlySerSerThr 44
Db 238 GTGACCAAGTATGCTCATATCTTACGGGAACGCTAGACAGATGAGCCGAGGCACTGAA 297
```

```
Qy 45 LeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAsp 64
Db 298 ---GCATCAGAGCTCCCGCTTCCCGCTCCGCTTATATCAAAAGACAGAGCGGAGCAG 354
Qy 65 AspLeuLeuAspThrTTPAsnSerThrPhe---SerProIleSerGluValAlaGlnThr 83
Db 355 GAGCTCAATGGCGGAGCCCAACTCTTACATTAAGAACTGTGATGGCAGCGCAAACTTTACT 414
Qy 84 SerPheMetGluAspIleArgProIleLeu-----PheProAsnAsn 97
Db 415 TCCATTCCACAAGACTTAAACGGCTTCTCTGGGACAAAGAAATTTTGGCCCGCGCA 474
Qy 98 Gln-----ThrLeuGlnPheCysAsnGlnThrProProHileuValGly 112
Db 475 CAGAAATCCACATCCGCCCTCTTCCCAACTGCACTGATCCGATCCCGTATGATGTGA 534
Qy 113 ProIleArgValPheLeuAspGluProAspPheLysThrLeuGlu---LysIleTyrPro 131
Db 535 CCCATCAGCGCCCAACAGACACTGAGTCACTGACATGATTAAGAGCGGAGCTTGAGACT 594
Qy 132 AspThrHisAlaGlyGlyHisGlyMetProLysAspCysValAlaArgHisArgValAla 151
Db 595 CTTTGGCCCTGTGGTGGCCCTTCCAGCCTGAAATCTGCAATGCCCAGCATCAGTGGCT 654
Qy 152 IleIleValProTyrArgAspArgGluAlaHisIleuArgIleMetLeuHisAsnLeuHis 171
Db 655 ATTTGTGGCTTCTCCGAGATGATAGCCCATCTTATTACTTTCTCTCGGCAACATCCAC 714
Qy 172 SerLeuLeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsn 191
Db 715 CCATTGTGATGAACAGAGCATGCCATGCTATTCATTTGATGTAGAGCAGCAACAGGG 774
Qy 192 GlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyr 211
Db 775 AAGCCCTTAAATCGGGCTGGCCATGATGAACATGTGTATTGGAGCGCTTAAAGCTGATC 834
Qy 212 ProTTPGlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeu 231
Db 835 CAGTGGATGTTTATATATTCACATGTCCATCTTCTGCTTGGAGCAGCCGCAATCTT 894
Qy 232 TyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLys 251
Db 895 TACAACGTGCCAGCTCAGCGCGGACACATGTGAGTGTATAGACACGCTGAACCTTCAAG 954
Qy 252 LeuProTyrSerAlaIlePheGlyGlyLysSerAlaLeuThrLysAspHisIleuLysLys 271
Db 955 TTGCTTATCGATCAATATTGGAGGTGTTCGCAATGACCGGTGAGCACTTTCAGGCC 1014
Qy 272 IleAsnGlyPheSerAspAspPheTTPGlyTTPGlyGluAspAspAspLeuAlaThr 291
Db 1015 GTAAATGGGTTCTCAAACTCGTTCTTGGCTGGGCGGAGATGACACATGTCCAAAC 1074
Qy 292 ArgThrSerMetAlaGlyLeuLysValSerArgTyrProthGlnIleAlaArgTyrLys 311
Db 1075 AGCTTGAAGCAGCAAACTATTCAATCAAGTATCCGGTCAACATAGCTGCTAAGAG 1134
Qy 312 MetIleYHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrIleMetGly 331
Db 1135 ATGCTGAAGCATCAGAAAGAAAGCCCAATCTTAAG-----CGCTATATAAACTTA 1185
Qy 332 GlnGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrIleVal 351
Db 1186 CAGAAATGCGATGAGCAAAATAGAACAGATGAAATCAACTCATTAAGATTTCTATCTAC 1245
Qy 352 AsnLeuGluLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
Db 1246 AGCATCAAAACATTTCCAACTTCACTTGTACTTGAAGACAGACTA 1290

RESULT 4
US-09-745-763-84
; Sequence 84, Application US/09745763
; Patent No. US20020065394A1
```

[illegible]

```

QY      194  PheAsnArgGlyLysLeuMetAsnValAlGlyTyrAspValAlaSerArgLeuTyrProTir 213
Db      658  TTTTAAAGGGGCAAAACGTTGAACGTTGGGGGTGGCGAAGGCGCTCGTGTATGAAGAGTGG 717
QY      214  GlnGysPheIlePheHisAspValAspLeuLeuProGluAspAspAsnLeuTyrThr 233
Db      718  GACTGCGCTGTTCTTGACAGATGTTGACTCTCTGCGAATAATGACCAACATCTGTATGTG 777
QY      234  Cys--ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeu 252
Db      778  TGTACACCCCGGGGAGACCCCGCCAGTGGCCGTTGCTCTATGAACAAGTTTGATACAGCCTC 837
QY      253  ProTyrSerAlaIlePheGlyValLysSerAlaLeuThrLysAspHisLeuLysLysIle 272
Db      838  CCGTACACCCCAAGTACTCTGGAGAAGCTCTACGACTTACTCTCTGACCACTGACTGAAGATG 897
QY      273  AsnGlyPheSerAsnAspPheTyrGlyTyrGlyGluAspAspLeuAlaThrArg 292
Db      898  AATGGCTTCCCCCATGATGAACTATGGGGCGTGGGTGTGATGACGACTT  GCTAACAGG 956
QY      293  ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet 312
Db      957  GTGGCGCTCGGTGGGAAGATGATCTCCGGCCCCCATCTGTAGGACACTTAAGAATG 1016
QY      313  IleLysHisSerThrGluAlaThrAspProValAsnLysCysArgTyrLysIleMetGly 332
Db      1017  GTGAAGACCCGAGAGAGATGAAGGCAATGAAGAAATCCCAACAGATTGACTCTGTGCT 1076
QY      333  GlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsn 352
Db      1077  CGTACCCCAAGATTCCTGGACCGAAGATGGAGTGAATCATCAGATCATACAGTTGCTGCT 1136
QY      353  LeuGluLeuLysProLeuTyrThrArgAlaValAspLeu 366
Db      1137  CGAAGCTGGGGCCTCTTTATACCAACATCACAGACGACTT 1178

RESULT 5
US-11-097-143-18553/c
/ Sequence 18553, Application US/11097143
/ Publication No. US20050208558A1
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ APPLICANT: et al.
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ TITLE OF INVENTION: DROSOPHILA GENES.
/ FILE REFERENCE: C1000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 18553
/ LENGTH: 4093
/ TYPE: DNA
/ ORGANISM: DROSOPHILA
US-11-097-143-18553

```

Alignment Scores:

Pred. No.: 6,636-77 Length: 4093
Score: 675.00 Matches: 161
Percent Similarity: 50.00% Conservative: 59
Best Local Similarity: 36.59% Mismatches: 137
Query Match: 33.23% Indels: 83
DB: 10 Gaps: 11

US-10-661-430-1 (1-383) x US-11-097-143-18553 (1-4093)

```
QY 5 HisLeuAlaValAla-ArgLeuLysSerLeuLeuValLeuGlyValAlaLeuLeuVal 24
DB 2852 CATATTGGCTTGGAGAGGTATCTTATCAATTATTTTGGACCGATATTAGCACTT-- 2795
QY 24 HisAlaMetIleTyrlsIleProSerLeuTyrlsAlaLeuThr-----11661 41
DB 2794 ----CGGCGCGTGCACAGATATGCTCATATCTACGGGAACGCTAGCAGCGATGAGCCGG 2739
QY 41 YSerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSe 61
DB 2738 AGGCACTGAA---GCATCCAGGCTGCCGCTCCCGCGCTCTTATCAAAGACAGACA 2682
QY 61 rThiSerAspLeuLeuAspThrTrpAsnSerThrPhe---SerProIleSerGluVal 80
DB 2681 GCGGAGCAGAGCTCAATGGCGGACCCAACTCTACATTAAGACATGTGATGGCCAGCGC 2622
QY 80 lAsnGlnThrSerPheMetGluAspIleArgProIleLeu-----Ph 94
DB 2621 AAACCTTACTCTTCACTTCCACAAAGACTTAAACGCGCTTCCGCGGACAAAGAAATTTT 2562
QY 94 eProAspAsnGln-----ThiLeuGlnPheCysAsnGln-----105
DB 2561 GCGCCCGGACAGAAATCCATCCGCCCTCTTGGCACTGACATGCCATCCGATCCCGC 2502
QY 106 -----ThiProIleHisLe 110
DB 2501 TGATGGTAAAGCTCGGCTTGGCTCTTTCTTAGACAGATTAAATACGTCACCTGTGTT 2442
QY 110 uValGlyProIleArgValPheLeuAspGluProAspPheLysThrLeuGln--Lys11 129
DB 2441 AGGTGAGCCCATCCAGCCCAACACGACACTGTGACGTGATATGAGCGGAGCT 2382
QY 129 eTyrrProAspThrHisAlaGlyLysIleGlyMetCProLysAspCysValAlaArgHisAr 149
DB 2381 TGGAACCTCTTTGGCCGCTGGTGGGCGCTTCGAGCTGAAACCTGCAATGCCAGATCA 2322
QY 149 gValAlaIleIleValProTyrrArgAspArgGluAlaHisLeuArgIleMetLeuHisAr 169
DB 2321 CGTGGCTATGTGTGGCCCTTCGGGATGATACGCCCATCTATTACTTTCTCGGCNA 2262
QY 169 nLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrrAlaIlePheIleValGluGlnVal 189
DB 2261 CATCCACCATTTCTGATGAAACAGAGCATGCCCTATCCATTTTCATTGTGAGACGAC 2202
QY 189 lAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyrrAspValAlaSerAr 209
DB 2201 CAACGGGAAGCCCTTATCGGCTGCCATGATGAACATGTGTATTGGAGGCTTAA 2142
QY 209 gLeuTyrrProIrrGlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspAr 229
DB 2141 GCTTGACAGGTGGAGATGTTTATATATCCAGATGTCGATCTTGGCTTGGAGACGCG 2082
QY 229 gAsnLeuTyrrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPheAs 249
DB 2081 CATCTTTTAACTGTCACGTCAGCGGACACATGTCTAGTGGCTATGAGACACGCTGAA 2022
QY 249 n-----Ty 250
DB 2021 CTTTCAGGCAAGTAGCAAGATGAACAATATCAATGTAATGATTCATTCATTTTA 1962
QY 250 rLysLeuProTyrrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuTy 270
```

```
DB 1961 TAGGTTCCTTATCGATCAATATTTGGAGGTGTTCCGCAATGACCGTGAACCTTCA 1902
QY 270 slYsIleAsnGlyPheSerAsnAspPheTrpGlyTrpGlyGlyLysAspAspLeuAl 290
DB 1901 GGGCTTAATGGGTTCTAAACTGTTTGGCTGGGCGGCGAGATGAGACATGTC 1842
QY 290 aThrArgThrSerMetAlaGlyLeuLysValSerArgTyrrProThrGlnIleAlaArgTy 310
DB 1841 CAACAGGTGAAGACGCAACGCAATCTATTCATATCAAGGATCCGGTCAACATGCTCGCTA 1782
QY 310 rLysMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrls11 330
DB 1781 CAAGATGCTGAAGCATCAAGAAAGCCAAATCTTAAGCGGTGATGACATCAATCAT 1722
QY 330 emet----- 331
DB 1721 TATTCATATTTTGAACCTAAATATGGAATTTTACCTTCTTTTGGACAGCTATGA 1662
QY 332 -----GlyGlnThrLysArgArgTrpThrArgAspGlyLeuSerAsnLeuTy 347
DB 1661 AAACCTTACGAATGGCATGAGCAAAATAGAA-----CAGATGGAATCACTCGATTA 1608
QY 347 sTyrlsLeuValAsnLeuGlyLeuLysProLeuTyrrThrArgAlaValAspLeu 366
DB 1607 GTATTCTATCTACAGCATCAAAACATTTCCAACTTCACTTGATCTTAGCAGACTA 1550
RESULT 6
US-10-087-192-2048
/ Sequence 2048, Application US/10087192
/ Publication No. US20020182586A1
/ GENERAL INFORMATION:
/ APPLICANT: Engelhard, David W.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
/ FILE REFERENCE: 529452000122
/ CURRENT APPLICATION NUMBER: US/10/087,192
/ PRIOR FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ NUMBER OF SEQ ID NOS: 2059
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2048
/ LENGTH: 3953
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-10-087-192-2048
Alignment Scores:
Pred. No.: 6,136-76 Length: 3953
Score: 667.50 Matches: 123
Percent Similarity: 65.06% Conservative: 52
Best Local Similarity: 45.72% Mismatches: 93
Query Match: 32.87% Indels: 1
DB: 5 Gaps: 1
US-10-661-430-1 (1-383) x US-10-087-192-2048 (1-3953)
QY 99 ThrLeuGlnPheCysAsnGlnThrProIleAsnValGlyProIleArgVal--Phe 117
DB 551 TCGCTGCCAGCTTGGCCCTGAGAGTCCCGCGCTCGTGTGGGCCCATGCTGATTCAT 610
QY 118 LeuAspGluProAspPheLysThrLeuGlyLysIleTyrrProAspThrHisAlaGlyGly 137
DB 611 AATATTGCTGTGATCTGAGCTTTGGCAAGAAAGAACCCAGAGATAAAGACGGGCGGC 670
QY 138 HisGlyMetCProLysAspCysValAlaArgHisArgValAlaIleIleValProTyrrArg 157
DB 671 CGTATCTCCCAAGAGACTGTGTCTCTCTCAAAAGTGGCGATGATCATCCATTCGCT 730
QY 158 AspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGln 177
```

Db 731 AACCGCAGAGATCTCAAAATGCTGCTGATATTGATTCCTTCAGCCGAG 790
QY 178 GlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGly 197
Db 791 CAACCTGACCTATGCGCATCTACGTCATCAACAGCTGAGACACCATGTTCAATCCAGCT 850
QY 198 LysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProThrGlnCysPheIle 217
Db 851 AAGCTCTCATATTTGGCTTTCAAGAGCCCTTGAAGACATATGATTAACAACCTCTTGG 910
QY 218 PheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGln 237
Db 911 TTCAGAGATGAGACCTCATTCGATGAGACGACGTAATGCTACAGGTGTTTGGCAG 970
QY 238 ProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIle 257
Db 971 CCACGGCACATTTCTGTGGCAATGACAAAGTTGGGTTTACCTGCCATATGTTCAAT 1030
QY 258 PheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsn 277
Db 1031 TTGGAGAGTCTCTGCTCTCAGTAACAAACAGTTCTTCCATTCATGAGATTCCTTAAT 1090
QY 278 AspPheTyrGlyTyrGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGly 297
Db 1091 AATTATTTGGGCTGGGAGAGAAAGATGACGACATTTTAAAGATTAGTTCAATAAGC 1150
QY 298 LeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThr 317
Db 1151 ATCTCATATCAACGTCCTCAAAATGCTGTAAGGAGGTGCAATGATCCGCAATTCAG 1210
QY 318 GluAlaThrAsnProValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArg 337
Db 1211 GACAAAGAAAATAGCCCAATCTCTCAGAGCTTGCAGCGGATCCGACATCAAGAGAAAC 1270
QY 338 TrpThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysPro 357
Db 1271 ATGCGGTGATGTTGGAATCACTTACCTACAGAGTGTGATGACAGATACCCG 1330
QY 358 LeuTyrThrArgAlaValValAspLeu 366
Db 1331 TTATATATCCCAATCAACAGTGACATC 1357

RESULT 7
US-10-466-941-5
; Sequence 5, Application US/10466941
; Publication No. US20040214273A1
; GENERAL INFORMATION:
; APPLICANT: Tatsui, Seki
; APPLICANT: Kazuhito, Fujiyama
; TITLE OF INVENTION: Methods for excretory production of glycoproteins
; TITLE OF INVENTION: having human-type sugar chains
; FILE REFERENCE: J198020401
; CURRENT APPLICATION NUMBER: US/10/466, 941
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: JP200112519
; PRIOR FILING DATE: 2001-01-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
US-10-466-941-5

Alignment Scores:
Pred. No.: 1.24e-76 Length: 1158
Score: 666.00 Matches: 137
Percent Similarity: 54.69% Conservative: 73
Best Local Similarity: 35.68% Mismatches: 128

Query Match: 32.79% Indels: 46
DB: 8 Gaps: 5
US-10-661-430-1 (1-383) x US-10-466-941-5 (1-1158)
QY 14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleTyr----- 29
Db 37 CTGTCGCGCGTCTGCGCTGCACTTGAGCGTCAACCTGCTTATTAACCTGCGTGGCCG 96
QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIle 46
Db 97 GACCTGAGCGCGCTGCCCACTG-----GTGGAGTCTTCACACCGCTG 141
QY 47 AlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
Db 142 CAGGCGGCTCGAACAAGTCCCGCCATCGGAGCTCTCCGGAGAGCTCCGACCGGA 201
QY 61 -----SerThrSer 63
Db 202 GGGGCGCGCGCGCTCTAGGCGCTCTCCACGCGCGCGGAGTGGGCACTCC 261
QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
Db 262 AGCCCACTGCTGATTCGCGCTGAGCGCGCTGAGCACTGACCTCGGTC----- 312
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
Db 313 -----CCAGTCCCCACACACACGCACTGCTGCCGCTGC 351
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
Db 352 CTGAGAGAGTCCCGCTACTAGTGGGCGCCAGCTGATTAACATGCTGAGAC 411
QY 123 PheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyHisGlyMetProLys 142
Db 412 CTGCACTGCTGCGCAAGAGAACCCCAATGTGAAGTGGCGCGCTATGCGCCAGG 471
QY 143 AspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
Db 472 GACTGCTCTCTCTCTCAAGAGTGGCCATCATTCATTCGCGCAACCGGACGAGAC 511
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnLeuAspTyrAla 182
Db 532 CTCAGTACTGCTATATTATTGTCACCCAGCTCTGAGCGCAGCAGCTGACTATGCG 591
QY 183 IlePheIleValGluGlnValAlaAsnGlnThrPheAspArgGlyLysLeuMetAsnVal 202
Db 592 ATCTATGTTATCAACACGCGGAGACACTATATTCATATGCTCAAGCTCTCAATGTT 651
QY 203 GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAsp 222
Db 652 GGCCTTCAGAGACCTTGAAGAGCTATGACTACACCTGCTTGTGTTAGTACGCTGAC 711
QY 223 LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
Db 712 CTCATTCAATGATGATCACTAATGCGTACAGGTGTTTTCACAGCCAGCGCACATTTCC 771
QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyLysIleSer 262
Db 772 GTTGCAATGATTAAGTTTGATTCAGCTTCACTTATGTTAGATTGAGAGTCTCT 821
QY 263 AlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAspAspPheTyrGlyTyr 282
Db 832 GCTCTAAGTAAACAACAGTTTCAACCATCAATGATTCCTTAATAATTATGGGCTGG 891
QY 283 GlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
Db 892 GAGAGAGAGATGATGATTTTAAACAGATTATTTAGAGGCACTGCTATATCTCGC 951
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
Db 952 CCAATGCTGTGTCGGAGAGTGTCCATGATCCGCACTCAAGAGACAAAGAAATGAA 1011

```
QY 323 ValAsnLysCysArgTyrIleMetGlyInThrIleArgTrpThrArgAspGly 342
    |||
    |||
    |||
Db 1012 CCCAATCTCCAGAGGTTTACCGAATTGCACACAAAGAGACATGCTCTGATGCT 1071
QY 343 LeuSerAsnLeuLysTyrIleValAsnLeuGlyLeuLysProLeuTyrThrArgAla 362
    |||
    |||
    |||
Db 1072 TTGAACCTCACTCACCTACCGAGTGTGATGATACAGAGATACCATTTGATATCCCAATC 1131
QY 363 ValValAspLeu 366
    |||
    |||
    |||
Db 1132 ACAGTGACATC 1143

RESULT 8
US-10-870-635-5
; Sequence 5, Application US/10870635
; Publication No. US20050143564A1
; GENERAL INFORMATION:
; APPLICANT: Tatsuji, Seki and Kazuhito, Fujiyama
; TITLE OF INVENTION: Methods for manufacturing of glycoproteins
; TITLE OF INVENTION: having human-type glycosylation
; FILE REFERENCE: 62, 736B
; CURRENT APPLICATION NUMBER: US/10/870,635
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 09/857,651
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1158
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
US-10-870-635-5

Alignment Scores:
Pred. No.: 1,24e-76 Length: 1158
Score: 666.00 Matches: 137
Percent Similarity: 54.69% Conservative: 73
Best Local Similarity: 35.68% Mismatches: 128
Query Match: 32.79% Indels: 46
DB: 9 Gaps: 5

US-10-661-430-1 (1-383) x US-10-870-635-5 (1-1158)
QY 14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleTyr----- 29
    |||
    |||
    |||
Db 37 CTGCTGGCGCTGCGCTGCACCTTGCGCTCACCTGTTTACTGCTGGCGCGC 96
QY 30 -----LysIleProSerLeuTyrGluAmnLeuThrIleGlySerSerThrIleuIle 46
    |||
    |||
    |||
Db 97 GACCTGAGCGCGCTGCCCACTG-----GTGCGAGTCTCCACACCGCTG 141
QY 47 AlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
    |||
    |||
    |||
Db 142 CAGGCGCGCTCAACAGTCCGCCCGCATCGGAGTCTCTCGGAGAGCTCCGAGCCGA 201
QY 61 -----SerThrSer 63
    |||
    |||
    |||
Db 202 GGGGCGCGCGCGCTCTAGGCGCGCTCTCCACGCGCGCGCGGGGTGGCGACTCC 261
QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
    |||
    |||
    |||
Db 262 AGCCAGTGTGATTTCTGCGCTGCGCTGAGGAACTTAACTCTGGCTC----- 312
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
    |||
    |||
    |||
Db 313 -----CCAGTGGCCCAACACACGACCGGACTGTCGCTGCCCGCTGC 351
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
    |||
    |||
    |||
Db 352 CCTGAGAGTCCCTCACTAGTGGGCCCATGCTGATTTGAATTTAAACATGCTGTGGAC 411
```

```
QY 123 PheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyValHisGlyMetProLys 142
    |||
    |||
    |||
Db 412 CTGGAGCTCTGGGCAAGCAGAACCCAAATGCAAGATGGGCGCGCTATCCCTCCAGG 471
QY 143 AspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
    |||
    |||
    |||
Db 472 GACTGCTCTCTCCACAGAGTGGCATTCATTCATTCCTCCGACCGGAGGAGGAC 531
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrAla 182
    |||
    |||
    |||
Db 532 CTCAGTACTGGCTATATTTATTTGACCCAGCTCGACCGCCGACACCTGACATATGAGC 591
QY 163 IlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnVal 202
    |||
    |||
    |||
Db 592 ATCTATGTTATCAACAGCGGAGACACTATATCAATCGGTCTAGCTCTCTCAATGTT 651
QY 203 GlyTyrAspValAlaSerArgLeuTyrProTrpGlnCysPheIlePheHisAspValAsp 222
    |||
    |||
    |||
Db 652 GGCTTCAAGAAAGCTTGAAGACTATGACTACACTGCTTTGTGTTAGTACCGTGAC 711
QY 223 LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
    |||
    |||
    |||
Db 712 CTCATTCCAATGAAATGACCATATGCTGACAGTGTGTTTTCACACGCGACACATTTCC 771
QY 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyValIleSer 262
    |||
    |||
    |||
Db 772 GTTGCAATGATTAATGTTGATTCAGCTTACCTTATGTTCACTATTTTGGAGGTGCTT 831
QY 263 AlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTrpGlyTyr 282
    |||
    |||
    |||
Db 832 GCTTAAGTAAACACACAGTTCTTACCATCATGATGATTTCCATATATATTTGGGGCTGG 891
QY 283 GlyIleGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
    |||
    |||
    |||
Db 892 GGAGGAGAAATGATGACATTTTAAACAGATTAGTTTAAAGGATGATGATATCTCGC 951
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
    |||
    |||
    |||
Db 952 CCAATGCTGCTGGTGGAGGTGCTGATGATCCCGCTCAAGAGACAAAGAAATGAA 1011
QY 323 ValAsnLysCysArgTyrIleValIleMetGlyInThrIleArgTrpThrArgAspGly 342
    |||
    |||
    |||
Db 1012 CCCAATCTCCAGAGGTTTACCGAATTGCACACAAAGAGACATGCTCTGATGCT 1071
QY 343 LeuSerAsnLeuLysTyrIleValAsnLeuGlyLeuLysProLeuTyrThrArgAla 362
    |||
    |||
    |||
Db 1072 TTGAACCTCACTCACCTACCGAGTGTGATGATACAGAGATACCATTTGATATCCCAATC 1131
QY 363 ValValAspLeu 366
    |||
    |||
    |||
Db 1132 ACAGTGACATC 1143

RESULT 9
US-10-084-817-168
; Sequence 168, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Pilon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; PRIOR APPLICATION NUMBER: 2002-02-25
; PRIOR FILING DATE: 60/270,784
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 168
; LENGTH: 4162
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119099A1 5549580CB1
US-10-084-817-168

Alignment Scores:
Pred. No.: 1,06e-75 Length: 4162
Score: 666.00 Matches: 137
Percent Similarity: 54.69% Conservative: 73
Best Local Similarity: 35.68% Mismatches: 128
Query Match: 32.79% Indels: 46
Gaps: 5

US-10-661-430-1 (1-383) x US-10-084-817-168 (1-4162)
QY 14 LeuLeuValLeuCySaValaValLeuLeuValHisAlaMetIleTyr----- 29
Db 221 CTGCTGGCGCTGCGCTGCGCACTTGCGTCACTTCCTTACTTACTGCTGCGCGC 280
QY 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIle 46
Db 281 GACCTGAGCGCGCTGCCCACTG-----GTGGAGTCTCACAACCGCTG 325
QY 47 AlaAspValaAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
Db 326 CAGGGCGCGCTCGAACAAGTCCGCCCATGGGAGTCTCCGGGAGCTCCGGACCGCA 385
QY 61 -----SerThrSer 63
Db 386 GGGGCGCGCGCGCTCTCTAGGCGCTCTCCAGCGCGCGCGGTGGCGACTCC 445
QY 64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValaAsnGlnThr 83
Db 446 AGCCCGAGTCTGGATTCTGGCCCTGGCCCGCTAGCACTTGACTTGCTGCTGCTC----- 496
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
Db 497 -----CAGTGCCTCCACCAACCAACCGACCTGCGTGGCGCGCTGC 535
QY 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
Db 536 CCGTGGAGTCCCGCTGCTTGTTGGGCGCCCATGCTGATTGATTAACTAGCTGTGGAC 595
QY 123 PheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyHisGlyMetProLys 142
Db 596 CTGGAGCTCGTGGCAAGCAAGACCCAAATGTGAAGTGGCGCGCTATGCCCAAG 655
QY 143 AspCyValaAlaArgHisArgValaAlaIleIleValProTyrArgAspArgGluAlaHis 162
Db 656 GACTGGCTCTCTCTCAACAAGTGGCCATCATTCATTCGCGCAACCGGCAAGACAC 715
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrAla 182
Db 716 CTCAAGTACTGGCTATATTATTTCACCCAGTCTCGAGCGCCGACGCTGACTATGGC 775
QY 183 IlePheIleValGluGlnValaAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnVal 202
Db 776 ATCTATGTTATCAACAGCGCGGAGACACTATATTCACTGCTGTAAGCTCTCAATGTT 835
QY 203 GlyTyrAspValaAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValaAsp 222
Db 836 GCGCTTCAAGAAGCTTGAAGACTATGACTACACCTCTTGTGTTAGTAGAGTGGAC 895
QY 223 LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
Db 896 CTCATTCCAATGATGACCATATATGCGTACAGGTGTTTTCACAGCCACCGCAATTTCC 955
QY 243 ValaIleAlaAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyIleSer 262
Db 956 GTTGCAATGGATTAGATTGATGACCTTACCTTATGTTCAAGTATTTGGAGGTGCTCT 1015
QY 263 AlaLeuThrLysAspHisLeuLysIleValaAsnGlnPheSerAsnAspPheTrpGlyTrp 282

```

```

Db 1016 GCTCTAAGTAAACAACAGTTTCTTAACCATGATGATTTCTTAATATATATGGGCTGG 1075
QY 283 GlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
Db 1076 GGAGGAGAAAGATGATGACATTTTAAACAGATTAGTTTAAAGGCGATGCTATATCTCC 1135
QY 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
Db 1136 CCAATGCTGTGTGTCGGAGGTGTGCGATGATTCGCACTCAAGAGCAAGAAAATGAA 1195
QY 323 ValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgArgTrpThrArgAspGly 342
Db 1196 CCCAATCTCAGAGGTTTACCGAATTCGACACACAAAGAGACATGCTCTTGATGTGT 1255
QY 343 LeuSerAsnLeuLysTyrLysLeuValaAsnGluLeuLysProLeuTyrThrArgAla 362
Db 1256 TTGAATCTCACTCACTTACCAAGGTGCTGATGTACAGATACCATTGTATACCAATTC 1315
QY 363 ValValaAspLeu 366
Db 1316 ACAGTGGACATC 1327

RESULT 10
US-10-761-435A-19
; Sequence 19, Application US/10761435A
; Publication No. US2004024181.7A1
; GENERAL INFORMATION:
; APPLICANT: Unema, Pablo
; APPLICANT: Bruenker, Peter
; APPLICANT: Ferreira, Claudia
; APPLICANT: Suter, Tobias
; TITLE OF INVENTION: Fusion Constructs and Use of Same to Produce Antibodies with
; FILE REFERENCE: 1975.0180003
; CURRENT APPLICATION NUMBER: US/10761,435A
; PRIOR FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US 60/441,307
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/491,254
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/495,142
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Nucleotide sequence of ManII-GalT
US-10-761-435A-19

Alignment Scores:
Pred. No.: 2.49e-76 Length: 1116
Score: 663.50 Matches: 133
Percent Similarity: 56.52% Conservative: 75
Best Local Similarity: 36.14% Mismatches: 129
Query Match: 32.67% Indels: 31
Gaps: 6

US-10-661-430-1 (1-383) x US-10-761-435A-19 (1-1116)
QY 16 ValLeuCySaValaValLeuLeuValHisAlaMetIle----- 28
Db 40 ATCTTCTGTGTGGATTCTTCTGCTTACTGATGCTGAGACGGGCTCACTTAGACTAC 99
QY 29 TyrLysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIleAlaAsp 48
Db 100 CCGAGAACCGCGCGCGGCGCTCTTCCATGAGCGCAAGCTTCAATAGTTGCAAGAA 159
QY 49 --ValaAspAlaMetGluAlaValaLeuGlyAsnThrAlaSerThrSerAspLeuLeu 67

```

Db 160 AAAATAGACATTGGAGCGTTGGTACCTGAGAAATAGATCATCTCAATAATTAGA 219
Qy AspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPheMetGlu 87
Db 220 GAC-----TCAGTCATCATATTGAGTAGTCTGTGGAG 252
Qy 88 AspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPhe----- 102
Db 253 GAT-----GATCGGAAAAGTTTCAAAAGCAATTTTCAGCCAGAGTGTGGC 297
Qy 103 -----CysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAsp 119
Db 298 TCACCCGCGCTCCGCTGAGAGAGTCCCGCTGTGGGCGCCATGCTGATTGATTAAAC 357
Qy 120 GluPro---AspPheSerThrLeuGluValIleTyrProAspThrHisAlaGlyGlnHis 138
Db 358 ATGCGCTGTGAGCTGAGCTCGTGGCAAGCAGAACCCAAAGTGAAGATGGAGCGGCCCC 417
Qy 139 GlyMetProIleAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAsp 158
Db 418 TATGCCGCCAGGAGACTGCGCTCTCTCTCAAGAGTGGCCATCATCTTCATTCGCCGAC 477
Qy 159 ArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaIleGlnGln 178
Db 478 CGCGCAGAGCAGCCTCAAGTACTGCTATATTGACCCAGTCTCGACGCGCAGCAG 537
Qy 179 LeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLys 198
Db 538 CTGAGACTATGGACTATATGTATCAACAGCGGGAGACACTATATCAATGATGCTTAAG 597
Qy 199 LeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProIleGlnCysPheIlePhe 218
Db 598 CTCCTCAAGTGTGGCTTCAAGAGCTTGAAGAGACTAGACTACACCTGTTGCTTT 657
Qy 219 HisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnPro 238
Db 658 AGTGAAGTGAAGCTCATCTCAATGAATGACCATATGCGTACAGGTTTTCACAGCCA 717
Qy 239 ArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePhe 258
Db 718 CGGCACTATTCCGTCGATGATGAATGATTGATGAGCTTACCTTATGTCAGTATTTT 777
Qy 259 GlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleAsnGlyPheSerAsnAsp 278
Db 778 GGAAGGTCTCTGCTCTAAGTAAACAAGCTTCTTACCATCAATGATTTCTTATATAT 837
Qy 279 PheTyrGlyTyrGlyGlyLysAspAspLeuAlaThrArgThrSerMetAlaGlyLeu 298
Db 838 TATGGGCGCTGGGAGAGAAATGATGACATTTTAAACAGATTAGTTTGAAGGCGATG 897
Qy 299 LysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGlu 318
Db 898 TCTATATCTCGCCAAATGCTGTGTGGGAGGTGTGCGCATGATCGCCACCAAGAGAC 957
Qy 319 AlaThrAspProValAsnLysCysArgTyrLysIleMetGlyIleThrLysArgArgTyr 338
Db 958 AAAAAAAGAACCCATCTCAGAGCTTTGACCCGAATTCACACACCAAGAGACAAAG 1017
Qy 339 ThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeu 358
Db 1018 CTCTCTGATGAGTTTGAATCACTACCTACAGGTCTGAGATGTAACAGATACCATTTG 1077
Qy 359 TyrThrArgAlaValAspLeu 366
Db 1078 TATACCCCAATCAACAGTGAATC 1101

APPLICANT: Berghs, Constance,
APPLICANT: Boldog, Ference,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomo R.,
APPLICANT: Eisen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gangoli, Esha A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: MaayanKar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Muradharu,
APPLICANT: Patnurejan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vernier, Corine A.M.,
APPLICANT: Zernusen, Bryan D.,
APPLICANT: Zhong, Mei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Number of SEQ ID NOS: 673
SOFTWARE: CuroseqList version 0.1
SEQ ID NO 81
LENGTH: 1119
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1116)
US-10-287-226-81
Alignment Scores:
Pred. No.: 5.35e-76 Length: 1119

Score: 661.00 Matches: 141
 Percent Similarity: 56.87% Conservative: 70
 Best Local Similarity: 38.01% Mismatches: 133
 Query Match: 32.55% Indels: 27
 DB: 7 Gaps: 7

US-10-661-430-1 (1-383) x US-10-287-226-81 (1-1119)

```

QY 7 AlAlValAlAArgLeuLysSerLeuValLeuCysAlaValLeuLeuValHisAl 26
DB 24 GCTGAGAGCGGCTGCAAGGCTGTCTCTCTCTGCTGCTGCACTTCTGCGCCGT 83
QY 26 aMetLeuTyLysIleProSerLeuTyGluAsnLeuThrIleGlySerSer---Thrl 45
DB 84 CATCTCTAC---TTTGACGCTTACGCGCCAGACCTGCGCTTCTGACGCGTTCAGTGC 140
QY 45 ulleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAlaSerThrSerAsp 65
DB 141 CCGAGGCGCTGCCAGCTCCCTCCAGCCAGCTGCTAGCAGAGCAGCAGCAGCACTG 200
QY 65 pLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPh 85
DB 201 CTCCCGCGCCAGCCAGCCAGCTCTAGCTCCGAGCTCCGAGTCCCGCAGTCC--- 225
QY 85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCysAsnG 105
DB 256 -----CTGCCCGGCTCCAGCGCTCCAGCTCCAGCTCCAGCTCCAGCTCCGA 296
QY 105 nThrProProHisLeuValGlyProIleArgValPheLeuAspGluProAspPheLeu 125
DB 297 CTGCGCACTGCTGTGTGGGAGACTGTGATCGAGTTCACCTCACC---ATGCC 350
QY 125 rLeuGluValIleTyLys-----ProAspThrHisAlaGlyGlyAsnGlyMetProLy 142
DB 351 CCTGAGAGCGGCTGCAAGGAGAAACCAGCGCTGCTAGCGCGCGCGAGAACACCGCC 410
QY 142 sAspCysValAlaArgHisArgValAlaIleIleValProTyArgAspArgGluAlaH 162
DB 411 CGACTGCACCCAGCCAGCAGAGCGGTGGCTGATCATCTCCCTTATGACACCGGAA 470
QY 162 sLeuAlaGlyIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnIleuAspTyAl 182
DB 471 CCGGCTGCTAGCTGCTCACTATCTACACCCCATCTTAGCGGAGCGGCTGCGCTAGCG 530
QY 182 aIlePheIleValGlyGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAspVa 202
DB 531 CGCTATGATGATCAACAGCATGCTGAGACACTTCAACCGCGCCCAAGCTCTTAACGT 590
QY 202 lGlyTyLys-----AspValAlaSerArgLeuTyLysProIleGlnCysPh 216
DB 591 GGGCTTCTAGAGCGCTGAGAGGAGATGCCGCC-----TATGACTGCTT 635
QY 216 eIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyLysProIle 236
DB 636 CATCTTAGAGATGTGACGCTGCTCCATGAGATGACCGCAACTTATACCGCTGCGGCA 695
QY 236 eGlnProArgHisMetSerValAlaIleAspLysPheAsnTyLysSerProTyLysSerAl 256
DB 696 CCAAGCCCGCCACTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
QY 256 aIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleAsnGlyPheSe 276
DB 756 CTACTTTGAGAGTGTGTCAAGCTGAGTAAAGCTCAGTTTCTGAGATCAATGAGCTTCCC 815
QY 276 rAsnAspPheThrGlyTyLysGlyGluAspAspArgLeuAlaThrArgThrSerMetAl 296
DB 816 CAATGAGTACTGCGGCTGCGGTGGCGAGATGATGATCTTCAACCGGATCTCCGAGC 875
QY 296 aGlyLeuLysValSerArgTyLysProThrGlnIleAlaArgTyLysMetIleLysHisSe 316
DB 876 TGGGATTAAGATCTCAGCGCCAGACATCCGAATTGGCGGCTGACCGCATGATCAAGCA 935
QY 316 rThrGluAlaThrAsnProValAsnLysCysArgTyLysIleMetGlyGlnThrLysAr 336

```

```

DB 936 CCGCGCAAGCATTAACGAACCTAACCCCTGAGAGTTTACCAAGATTCAAAACGAACT 995
QY 336 gArgTyPThrArgAspGlyLeuSerAsnLeuLysTyLysLeuValAsnLeuGlyLeuTy 356
DB 996 GACCATGAAGCGGAGCATTTGGGTGAGTGGGATCGAGTCTTGAGAGTGTCTGGCA 1055
QY 356 sProLeuTyThrArgAlaValAspLeu 366
DB 1056 ACCACTTTCACCAATATCAAGTGAAT 1086

```

RESULT 12

```

US-10-713-970-20
; Sequence 20, Application US/10713970
; Publication No. US20050106664A1
; GENERAL INFORMATION:
; APPLICANT: Contreras, Roland
; APPLICANT: Calleja, Nico L. M.
; APPLICANT: Vervecken, Wouter
; APPLICANT: Kaigorodov, Vladimir
; TITLE OF INVENTION: MODIFICATION OF PROTEIN GLYCOSYLATION IN METHYLOTROPHIC YEAST
; FILE REFERENCE: 17106
; CURRENT APPLICATION NUMBER: US/10/713, 970
; CURRENT FILING DATE: 2003-11-14
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-713-970-20

```

Alignment Scores:

Pred. No.:	6, 92e-76	Length:	1191
Score:	660.50	Matches:	138
Percent Similarity:	53.83%	Conservative:	73
Best Local Similarity:	35.20%	Mismatches:	134
Query Match:	32.52%	Indels:	47
DB:	9	Gaps:	5

US-10-661-430-1 (1-383) x US-10-713-970-20 (1-1191)

```

QY 7 AlAlValAlAArgLeuLysSerLeuValLeuCysAlaValLeuLeuValHisAl 26
DB 49 TCCCTACAGCGGCGCTGCGCGCTGCTGCGCTGCTGCGCTGCGCACTTGGCGCTCACC 108
QY 27 MetIleTyLys-----LysIleProSerLeuTyGluAsnLeu 38
DB 109 CTGCTTACTACTGCTGCTGCTGCGCGCACTGAGCGGCTGCTGCTGCTGCTGCTGCTG 156
QY 39 ThrIleGlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsn 58
DB 157 ---GTGGAGTCTCCACACCGCGTGCAGAGGCGGCTCGAAGGTGCGCGCGCATCGGGCAG 213
QY 59 ThrAla----- 60
DB 214 TCCCTCGGAGCTCGCGACCGGAGGCGCGCGCGCGCGCTGCTGAGCGCTCTCC 273
QY 61 -----SerThrSerAspAspLeuAspThrTrpAsnSerThrPheSer 75
DB 274 CAGCGCGCGCGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 333
QY 76 ProIleSerGluValAsnGlnThrSerPheMetGluAspIleArgProIleLeuPhePro 95
DB 334 AACTTGAACCTCGGCTC-----CCAGTCCCGCACACC 363
QY 96 AspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArg 115
DB 364 ACCGACATGCGCTGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 423
QY 116 ValPheLeuAspGluPro---AspPheLysThrLeuGluLysValLeuTyLysProAspThrHis 134
DB 424 ATTGAGTTTAACATGCTGTGAGCTGAGCTGCTGCGAAGCAAGAACCCAAATGTGAG 483

```


[illegible]

```

Db      942 CAAGCATTAAGCACTTAACCTTACCAAGATTCAAAACAGCAAGCTGACCAT 1001
Qy      338 pthrraagpbglyleuseranleuylstyrylsleuvalasnlleugluleuysprole 358
Db      1002 GAAGCGGAGCGGATGGTGGTACAGGCTTGTGGAGTGTCTCGGCAACCACT 1061
Qy      358 uTyrrThArgAlaValAspLeu 366
Db      1062 CTTCAACCAATATCAGACTGACATT 1086

```

RESULT 14

```

US-11-105-796-33
; Sequence 33, Application US/11105796
; Publication No. US20050181437A1
; GENERAL INFORMATION:
; APPLICANT: CLAUSEN, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GUCCOSAMINE B-1,
; FILE REFERENCE: 04305/100521-US1
; CURRENT APPLICATION NUMBER: US/11/105,796
; PRIOR FILING DATE: 2005-04-13
; PRIOR APPLICATION NUMBER: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 1116
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-105-796-33

```

Alignment Scores:

```

Pred. No.: 1.33e-75 Length: 1116
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
Gaps: 6

```

US-10-661-430-1 (1-383) x US-11-105-796-33 (1-1116)

```

Qy      7 AlaValAlaArgLeu-LysSerLeuLeuValLeuCyAlaValLeuLeuValHisAl 26
Db      24 GCTGAGCGGCTGCGAAGCTGCTCTTCTGTGCTGTGACTTCCTCGTGGCGCT 83
Qy      26 aMetIleTyrlsleProSerLeuTyrlsLeuAsnLeuThrIleGlySerSer---ThrIle 45
Db      84 CACTCTCTAC---TTTGACGCTTACGCCAGCACTGCGCTTCTTCAAGCGCTTCACTGC 140
Qy      45 uIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrIleAspThrIleSerAspAs 65
Db      141 CCGAGCGCCCTGCGCCATCCCTCCACCCAGCTGCTGAGAGGAGCGAGCGAGCGAGCAACTG 200
Qy      65 pLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGluValAsnGlnThrSerPh 85
Db      201 CTCGGCGGCCAAGCCACCGCGCTAGCTCCGGGCTCCCTGAGGTCCCAAGTGC----- 255
Qy      85 eMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCyAsnGln 105
Db      256 -----CTGCCCGGCTCCACCGGCTCCCGCTCCCGCTGCTGCTCTCTGA 296
Qy      105 nThrProHisIleuValGlyProIleArgValPheLeuAspGluPro---AspPheIle 124
Db      297 CTCGGCACTGCTGCTTGTGGGCAAGCTGATGAGTTCACCTTCCACCATGCGCTGGA 356
Qy      124 sthLeuGluIuValIleTyProAspThrHisAlaGlyValHisGlyMetProIlyAspCy 144
Db      357 GCGGCTGCAAGGAGAAACGAGCGCTGCTATGGCGCGCGGAGAACACCGCCGCACTG 416
Qy      144 sValAlaArgHisArgValAlaIleIleValProTyrrArgAspArgGluAlaHisIleuAr 164

```

```

Db      417 CACCCCGCCAGACGGTGGCGGCTATCATCCCTTTAGACACCGGGAAACACCACCTGGC 476
Qy      164 gIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyrrAlaIlePh 184
Db      477 CTACTGGCTCCACTATCTACACCCCATCTTGAAGCGGCGGCGGCTGCGGCTCTA 536
Qy      184 eIleValGluGlnValAlaAsnGlnThrPheAsnArgIlyLysLeuMetAsnValGlyTy 204
Db      537 TGTCAATCAACAGCATGTGAGAGCACCTTCAACCGGGCCAAAGCTGCTTACCTGGGCTT 596
Qy      204 r-----AspValAlaSerArgLeuTyrrProIleGlnCyPheIlePh 218
Db      597 CTTAGAGCGCTGAGAGAGATCCGCC-----TATGACTCTTCACTT 641
Qy      218 eHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrrThCyProIleGlnPr 238
Db      642 CAGCGATGTGACCTGTGCTCCCATGATGACCGCAACCTATACCGGTGGCGGCAACCC 701
Qy      238 oArgHisMetSerValAlaIleAspIlyAspLeuTyrrLysLeuProTyrrSerAlaIlePh 258
Db      702 CCGCAGCTTTGCCATTTGCATGAGCAAGTTTGCTTCCGGCTTCCCTATGCTGGCTACTT 761
Qy      258 eGlyGlyIleSerAlaLeuThrIlyAspHisIleuLysIleAsnGlyPheSerAspAs 278
Db      762 TGGAGGTGTGTACAGCGCTGAGTAAGCTCAGTTTGAAGAATCAATGCTTCCCAATGA 821
Qy      278 pHetTPGlyTyrrPglGlyGlyLysAspAspLeuAlaThrArgThrSerMetAlaGlyLe 298
Db      822 GTACTGGGCTGGGCTGGCGAGATGATGACATCTTCAACCGGATCTCCCTGACTGGAT 881
Qy      298 uLysValSerArgTyrrProThrGlnIleAlaArgIlyLysMetIleLysHisSerThrGl 318
Db      882 GAAGATCTACAGCCCGCAGACATCCGAATCGCGCTACCGCATGATCAAGACACACCGCA 941
Qy      318 uAlaThrAspProValAsnLysCyAspArgTyrrLysIleMetGlyGlnThrIlyAspArgTr 338
Db      942 CAAGCATTAAGCACTTAACCTTACCAAGATTCAAAACAGCAAGCTGACCAT 1001
Qy      338 pthrraagpbglyleuseranleuylstyrylsleuvalasnlleugluleuysprole 358
Db      1002 GAAGCGGAGCGGATGGTGGTACAGGCTTGTGGAGTGTCTCGGCAACCACT 1061
Qy      358 uTyrrThArgAlaValAspLeu 366
Db      1062 CTTCAACCAATATCAGACTGACATT 1086

```

RESULT 15

```

US-10-956-157-1114
; Sequence 1114, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Myeth
; APPLICANT: Mounets, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1114
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1114

```

Alignment Scores:

```

Pred. No.: 3.21e-75 Length: 1890
Score: 658.00 Matches: 139
Percent Similarity: 56.91% Conservative: 71
Best Local Similarity: 37.67% Mismatches: 136
Query Match: 32.40% Indels: 23
Gaps: 6

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 9, 2006, 04:21:46 ; Search time 982 Seconds
(without alignments)
284.350 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRIHVARILKSLVLCAV.....VLLKCKRRLRDRPFCF 383

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4637633 seqs, 364532575 residues
Total number of hits satisfying chosen parameters: 9275266

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+.p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10661430/runat_06012006_123139_6385/app_query.fasta_1.519
-DB=Published_Applications_NA_New -OPMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bl0sum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=pco -NORM=ext -HEAPSIZE=500 -MITLEN=0
-MAXLEN=200000000 -USER=US10661430 @CGN 1.1.122 @runat_06012006_123139_6385
-NCPU=6 -ICPU=3 -NO MMAP -LARGESUBSET -NEG_SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Published Applications_NA_New.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	642.5	31.6	1964	7	US-11-102-240-41 Sequence 41, Appl 1
2	635.5	31.3	951	7	US-11-136-527-1055 Sequence 1055, Ap 1
3	635.5	31.3	951	7	US-11-136-527-5151 Sequence 5151, Ap 1
4	583	28.7	4646	6	US-10-955-054A-120 Sequence 120, App 1
5	331.5	16.3	2507	6	US-10-750-185-30398 Sequence 30398, A
6	331.5	16.3	2507	6	US-10-750-623-30398 Sequence 30398, A
7	313	15.4	16822	6	US-10-995-561-13425 Sequence 13425, A
8	313	15.4	18705	6	US-10-995-561-13401 Sequence 13401, A

9	163	8.0	1728	6	US-10-750-185-43601 Sequence 43601, A
10	163	8.0	1728	6	US-10-750-623-43601 Sequence 43601, A
11	154	7.6	201	6	US-10-995-561-22169 Sequence 22169, A
12	154	7.6	201	6	US-10-995-561-68468 Sequence 68468, A
13	136.5	6.7	3458	6	US-10-750-185-34973 Sequence 34973, A
14	136.5	6.7	3458	6	US-10-750-623-34973 Sequence 34973, A
15	109.5	5.4	3745	7	US-11-000-688-1051 Sequence 1051, Ap 1
16	100.5	4.9	3877	7	US-11-102-240-71 Sequence 71, Appl 1
17	87	4.3	90572	7	US-11-124-366A-2900 Sequence 2900, Ap 1
18	84	4.1	3148	6	US-10-793-626-4059 Sequence 4059, Ap 1
19	83	4.1	1521	7	US-11-134-241-30 Sequence 30, Appl 1
20	83	4.1	1812	7	US-11-134-241-42 Sequence 42, Appl 1
21	83	4.1	3896	7	US-11-134-241-44 Sequence 44, Appl 1
22	82.5	4.1	777	6	US-10-793-626-235 Sequence 235, A
23	82.5	4.1	1054	6	US-10-750-185-41517 Sequence 41517, A
24	82.5	4.1	1054	6	US-10-750-623-41517 Sequence 41517, A
25	82.5	4.1	3446	6	US-10-793-626-3343 Sequence 3343, Ap 1
26	82.5	4.1	3827	6	US-10-793-626-3598 Sequence 3598, Ap 1
27	82.5	4.1	3919	6	US-10-793-626-3814 Sequence 3814, Ap 1
28	82.5	4.1	57489	6	US-10-927-466-100 Sequence 100, App 1
29	82	4.0	1212	6	US-10-821-234-638 Sequence 638, App 1
30	82	4.0	1311	7	US-11-000-688-484 Sequence 484, App 1
31	82	4.0	4665	7	US-11-136-527-3309 Sequence 3309, Ap 1
32	81.5	4.0	3514	6	US-10-750-185-44589 Sequence 44589, A
33	81.5	4.0	3514	6	US-10-750-623-44589 Sequence 44589, A
34	80.5	4.0	2475	6	US-10-131-826A-195 Sequence 195, App 1
35	80.5	4.0	6243	6	US-10-821-234-788 Sequence 788, App 1
36	80	3.9	810	7	US-11-134-241-26 Sequence 26, Appl 1
37	80	3.9	1101	7	US-11-134-241-38 Sequence 38, Appl 1
38	80	3.9	3405	6	US-10-518-599-1 Sequence 1, Appl 1
39	80	3.9	3732	6	US-10-131-826A-71 Sequence 71, Appl 1
40	80	3.9	3900	7	US-11-197-721-42 Sequence 42, Appl 1
41	79.5	3.9	1288	6	US-10-750-185-54909 Sequence 54909, A
42	79.5	3.9	1288	6	US-10-750-623-54909 Sequence 54909, A
43	79	3.9	2526	6	US-10-750-185-33449 Sequence 33449, A
44	79	3.9	2526	6	US-10-750-623-33449 Sequence 33449, A
45	79	3.9	190882	7	US-11-121-086-69 Sequence 69, Appl 1

ALIGNMENTS

RESULT 1
US-11-102-240-41
Sequence 41, Application US/11102240
Publication No. US20050260647A1
GENERAL INFORMATION:
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESSION
FILE OF INVENTION: ESOPHAGEAL TUMOR
CURRENT APPLICATION NUMBER: P1320R1C106C
CURRENT FILING DATE: 2005-04-08
PRIOR APPLICATION NUMBER: US/11102,240
PRIOR FILING DATE: 2002-05-07
PRIOR APPLICATION NUMBER: 10/006867
PRIOR FILING DATE: 2001-12-06
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: 60/170262
PRIOR FILING DATE: 199-12-09
NUMBER OF SEQ ID NOS: 170
SEQ ID NO 41
LENGTH: 1964
TYPE: DNA
ORGANISM: Homo Sapien
US-11-102-240-41
Alignment Scores: 5.76e-69 Length: 1964
Pred. No.:

Score: 642.50 Matches: 143
 Percent Similarity: 53.37% Conservative: 55
 Best Local Similarity: 38.54% Mismatches: 124
 Query Match: 31.63% Indels: 49
 DB: 7 Gaps: 8

US-10-661-430-1 (1-383) x US-11-102-240-41 (1-1964)

```

QY 5 HisLeuAlaValAla---ArgLeuLysSerLeuLeuValLeuCysAlaValLeuLeu--- 22
DB 140 CACCTTCCACAAATCCGATTACTGCTGCTTGTGACTTGTGCTGACAGTGGTGGG 199
QY 23 -----LeuValHisAlaMetIleTyrLysIleProSerLeuTyrGlu 36
DB 200 TGGGCGACACAGTAAGTACTGTCGTGGGGGCC---ATTCAAGAGATTCTTAAGCAAGAG 256
QY 37 AsnLeuThrIleGlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeu 56
DB 257 TTCATGGCTAATTTCCATTAAGACCCCTCATTTTG-----GGGAAGGAAAAAAGCTTG 307
QY 57 GlyAsnThrAlaSerThrSerAspAspLeuLeuAspThrTrpAsnSerThrPheSerPro 76
DB 308 ACTAATGAACATCCACAGAAAGATGAAGCTTGACAC----- 346
QY 77 IleSerGluValAsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAsp 96
DB 346 ----- 346
QY 97 AsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArgVal 116
DB 347 -----TGTCTTCTGTGTCCTTCACTCAGAGGCCAGAGCAAGCTC 388
QY 117 PheLeuAspGluProAspPheLysThrLeuGluLysIleTyrProAspThrHis----- 134
DB 389 ATTTTC---AAACCAAGATCTC---ACTTTGAAAGAGTATACAGCAAGAAATCCCAAGTG 442
QY 135 AlaGlyGlyHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleLeuVal 154
DB 443 TCCAGAGGCCGGATGCGCCTCAGGAATGTAAAGCTTTTACAGAGGGTGGCATCTCGTT 502
QY 155 ProTyrArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeu 174
DB 503 CCCACCGGAAACAGAGAAACACTGATGTACTGTGGAACATCTGCATCCCTTCCTG 562
QY 175 AlaLysGlnLeuAspArgTyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPhe 194
DB 563 CAGAGCGACAGCTGGATTTGGCATCTACGTCTACACAGGTGTAAGGTAAGTTT 622
QY 195 AsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTrpGln 214
DB 623 AATCGAACCAAACTCTTGAATGTGGCTATCTAGAACCCCTCAAGGAAGAAATTGGGAC 682
QY 215 CysPheIleIleHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCys 234
DB 683 TGCCTTATATTTCCACGATGTGAGCTGTGATCCGACGATGACTTTAACCTTTTCAAGTGT 742
QY 235 ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyr 254
DB 743 GAGGAGCATCCCAAGACTCTGGTGTGGCAGAGAACGACACTGGGTACAGGTATCGTTAC 802
QY 255 SerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysIleAsnGly 274
DB 803 AGTCGATATTTTGGGGGTGTTACTGCCCTTAAGCAGAGAGGAGTTTCAAGGTGAATGGA 862
QY 275 PheSerAsnAspPheTrpGlyTyrGlyGlyGluAspAspLeuAlaThrArgThrSer 294
DB 863 TTCTCTAACAACTACTGGGATGGGAGGCCAAGACATGACCTCAGACTCAGAGTTGAG 922
QY 295 MetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMetIleLys 314
DB 923 CTCGAAGAATGAATTTTCCCGGCCCTGCTTAAGTGGGTAAATTATCAATGTTCTTC 982
QY 315 HisSerThrGluAlaThrAsnProValAsnLysCysArgTyrLysIleMetGlyGlnThr 334

```

```

DB 983 CACACTAGAGACAAGGCAATAGAGTGAACGCAAGACGATGAAGCTTTACACCAAGTG 1042
QY 335 LysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGlu 354
DB 1043 TCACGAGTCTGGAGAAACAGATGGGTGAGTAGTGTCTTATTAATTAATGATCTGGAA 1102
QY 355 LeuLysProLeuTyrThrArgAlaValAsp 365
DB 1103 CACAATCTTTATATATCAACATCAAGTGGAT 1135

```

RESULT 2

```

US-11-136-527-1055
; Sequence 1055, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Nounes, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1055
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; US-11-136-527-1055

```

Alignment Scores:

```

Pred. No.: 1,366-68 Length: 951
Score: 635.50 Matches: 117
Percent Similarity: 62.83% Conservative: 52
Best Local Similarity: 43.49% Mismatches: 99
Query Match: 31.29% Indels: 1
DB: 7 Gaps: 1

```

US-10-661-430-1 (1-383) x US-11-136-527-1055 (1-951)

```

QY 99 ThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArgValPheLeu 118
DB 133 ACTCTGCTGCTGCTGCTGCTGAGAGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 192
QY 119 AspGluPro---AspPheLysThrLeuGluLysIleTyrProAspThrHisAlaGly 137
DB 193 AATATTCCTGTGATCTGAGCTTTTGGCAAGAAAGAACCAAGATTAAGATGGGGG 252
QY 138 HisGlyMetProLysAspCysValAlaArgHisArgValAlaIleLeuValProTyrArg 157
DB 253 CGTTACTTCCCAAGGAGTATCTTCTCCCTCACAAGTGGCCATCTATCCATTCCT 312
QY 158 AspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGln 177
DB 313 AACCGCAGAGGACCTCAAAATCTGGCTGTATATTGATTCAGATCTTCACGCGCAG 372
QY 178 GlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAspArgGly 197
DB 373 CAATCGACTATGGCAATCTAGTCATCAATCAAGCTGGAGACCAATGTTATTCAGACT 432
QY 198 LysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTrpGlnCysPheIle 217
DB 433 AAGCTGCTCAAGCTTGGCTTTCAAGAGGCTTGAAGAAGTATGACTACAACTGTTTGG 492
QY 218 PheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGln 237
DB 493 TTCAGATGTGACCTCATTTCAATGATGACCATATGACCAATGACAGTGCCTTTACAG 552
QY 238 ProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIle 257
DB 553 CACGGCATATTTCTGTGCAATGACAGTGGGTTTACCTGCTTACGTTCAATAT 612

```

```

QY 258 PheGlyIleSerAlaLeuThrIysAspHisIleuLysIleAsnGlyPheSerAsn 277
DB 613 TTTGAGGTCGTCCGTCCTCAAGTAAACAAGTTCCTTACCATCAATGATTTCTTAAT 672
QY 278 AspPheTrpGlyTrpGlyGluAspAspLeuAlaThrArgThrSerMetAlaGly 297
DB 673 AATTACTGGCGCTGGGAGGAGAAAGATGATGACATTTTAAACAGATTGATTGATAAGGC 732
QY 298 LeuLysValSerArgTyrrProThrGlnIleAlaArgTyrrLysMetIleLysHisSerThr 317
DB 733 ATGCTATATATCAGCCCAATGCTGTGAGCGGTGCGATGATATCCGACATCAAGA 792
QY 318 GluAlaThrAspProValAsnLysCysArgTyrrLysIleMetGlyGlnThrIysArgArg 337
DB 793 GACAAAGTTTAATGAGCCCAACCTCAGAGGTTTGACCGGATGCACATACAAAGGAAC 852
QY 338 TrpThrArgAspGlyLeuSerAsnLeuLysTyrrLysIleValAsnLeuGluLeuLysPro 357
DB 853 ATGCCCGCTTGATGTTTGAACCTACCTTACTTACAGGTTGTGACATACAGAGATACCCG 912
QY 358 LeuTyrrThrArgAlaValAlaAspLeu 366
DB 913 TTATATACCAAAATACAGATGACATC 939

RESULT 3
US-11-136-527-5151
; Sequence 5151, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5151
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-5151

Alignment Scores:
Pred. No.: 1,366-68 Length: 951
Score: 635.50 Matches: 117
Percent Similarity: 62.83% Conservative: 52
Best Local Similarity: 43.49% Mismatches: 99
Query Match: 31.29% Indels: 1
DB: 7 Gaps: 1

US-10-661-430-1 (1-383) x US-11-136-527-5151 (1-951)
QY 99 ThrLeuGlnPheCysAsnGlnThrProProHisIleuValGlyProIleArgValPheLeu 118
DB 133 ACTGTCCTGCTGCTGCTGAGAGTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 192
QY 119 AspGluPro--AspPheLysThrLeuGluLysIleTyrrProAspThrHisIleGlyGly 137
DB 193 AATATATTCCTGTCGATCTGAGCTTTGGCAAGAAAGAACCCAKAGTAAAGATGGCGGS 252
QY 138 HisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIleValProTyrrArg 157
DB 253 CGTACTTCCCAAGGCTGATCTCCCTCACAAGGTGGCCATCATATCCATTCGCT 312
QY 158 AspArgGluAlaHisIleuArgIleMetLeuHisAsnLeuHisSerIleuLeuAlaLysGln 177
DB 313 AACCGGAGGAGACCTCAAAATACCTGCTGTATATTTGCTATCCAGTCTTCAGCGCCAG 372
QY 178 GlnLeuAspTyrrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGly 197

```

```

DB 373 CAACTCGACTATGCGATCTTACGTATCATAGCGCTGGAGACACCATGTTTATGAGCT 432
QY 198 LysLeuMetAsnValGlyTyrrAspValAlaSerArgLeuTyrrProTrpGlnCysPheIle 217
DB 433 AAGCTGCTCAACGTTGGCTTCAAGAGGCTTGAAGACTATGATGACAACTGCTTTTGG 492
QY 218 PheHisAspValAspLeuLeuProGluAspAspAsnLeuTyrrThrCysProIleGln 237
DB 493 TTCAGTATGAGGACTATTCATTCATGATGACATTAATGCCKACAGGCTTTTCACAG 552
QY 238 ProArgHisMetSerValAlaIleAspLysPheAsnTyrrLysLeuProTyrrSerAlaIle 257
DB 553 CCAGCGCATATTTCTGTGCAATGACAGATTCGGGTTTAGCTGCTTCATGATCAT 612
QY 258 PheGlyIleSerAlaLeuThrIysAspHisIleuLysIleAsnGlyPheSerAsn 277
DB 613 TTTGAGGTCGTCCGTCCTCAAGTAAACAAGTTCCTTACCATCAATGATTTCTTAAT 672
QY 278 AspPheTrpGlyTrpGlyGluAspAspLeuAlaThrArgThrSerMetAlaGly 297
DB 673 AATTACTGGCGCTGGGAGGAGAAAGATGATGACATTTTAAACAGATTGATTGATAAGGC 732
QY 298 LeuLysValSerArgTyrrProThrGlnIleAlaArgTyrrLysMetIleLysHisSerThr 317
DB 733 ATGCTATATATCAGCCCAATGCTGTGAGCGGTGCGATGATATCCGACATCAAGA 792
QY 318 GluAlaThrAspProValAsnLysCysArgTyrrLysIleMetGlyGlnThrIysArgArg 337
DB 793 GACAAAGTTTAATGAGCCCAACCTCAGAGGTTTGACCGGATGCACATACAAAGGAAC 852
QY 338 TrpThrArgAspGlyLeuSerAsnLeuLysTyrrLysIleValAsnLeuGluLeuLysPro 357
DB 853 ATGCCCGCTTGATGTTTGAACCTTACTTACCTTACAGGTTGTGACATACAGAGATACCCG 912
QY 358 LeuTyrrThrArgAlaValAlaAspLeu 366
DB 913 TTATATACCAAAATACAGATGACATC 939

RESULT 4
US-10-955-054A-120
; Sequence 120, Application US/10955054A
; Publication No. US20050266420A1
; GENERAL INFORMATION:
; APPLICANT: PUSZTAI, LAJOS
; APPLICANT: SYMMANS, W. FRASER
; APPLICANT: HESS, KENNETH R.
; APPLICANT: AYERS, MARK
; APPLICANT: STEC, JAMES
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
; FILE REFERENCE: UYXC:880US
; CURRENT APPLICATION NUMBER: US/10/955,054A
; PRIOR FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 120
; LENGTH: 4646
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-955-054A-120

Alignment Scores:
Pred. No.: 5,978-61 Length: 4646
Score: 583.00 Matches: 127
Percent Similarity: 54.22% Conservative: 53
Best Local Similarity: 38.25% Mismatches: 136
Query Match: 28.71% Indels: 16
DB: 6 Gaps: 7

US-10-661-430-1 (1-383) x US-10-955-054A-120 (1-4646)
QY 39 ThrIleGlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsn 58

```

[illegible]

```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30398
; LENGTH: 2507
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-30398

Alignment Scores:
Pred. No.: 5,31e-30 Length: 2507
Score: 331.50 Matches: 72
Best Similarity: 50.53% Conservative: 24
Percent Similarity: 37.89% Mismatches: 50
Query Match: 16.32% Indels: 44
DB: 6 Gaps: 3

US-10-661-430-1 (1-383) x US-10-750-185-30398 (1-2507)
QY 142 LysAspCySVaAlaIaRGHIsArGVaAlaIleIleValProLYrArGAspARgLuAla 161
DB 583 GAAGTGTCTTCTCTTCATCTCTTATTCCTCAACCTGCTTGACTTCCCTCCATTC 642
QY 162 HisLeuARgIleMeCLeuHIsAsnLeuHIsSerLeuEuaAlaYsGInLeuAspTYr 181
DB 643 ACTTTATGACTTCTCTTCAC-----GCCACTATAGACCTAGTTGG 684
QY 182 AlaIlePheIleValGluGInVaAlaAsnGInThrPheAsnARgGlyLYLeuMetAsn 201
DB 685 TCCTCCCTCCATTCATCTTCACAGCTGGAATGGAACATTTAACAGGGCAAACTGCTGAT 744
QY 202 ValGLYTYrAspVaAlaIsErArGLeuTYrProTgInCySPheIlePheHisAspVa 221
DB 745 GTCCGGGTGCCGGAGGCCCTGCCAGATGAAGAATGGAGCTGTTCTTGTGCATGATGT 804
QY 221 AspLeuLeuProGluAspAspArGAsnLeuTYrThrCys--ProIleGInProArGI 240
DB 805 GGATCTCTCGCAAGAGAACCAACAACCTGATATGTGCGATCCCGAGACCTCGGCA 864
QY 240 sMeCserVaAlaIleAspLYsPheAsnTYrLYs----- 251
DB 865 CTTTGCTGTGCATGAAACAATTGGATGACAGTAGAGGGCACAGAGGGTACTGGAAG 924
QY 251 ----- 251
DB 925 CAGGSTATTCACCATTCAGTGGAAAGAGGAGTGGCTCCCAAAAAGCTGTGAATC 984
QY 252 -----LeuProTYrSerAlaIlePheGlyIleSerAlaLe 264
DB 985 CAAGATCTTTCTTCCCGCAGCGCTCCGTAATCCCAAGTACTTTGGCGGGGTCTACGCCGT 1044
QY 264 uThrLYsAspHisLeuLYsLYsIleAsnGlyPheSerAsnAspPheTPrGlyTYrGLI 284
DB 1045 CACTCCTGACCAAGTACGTAAGATGAATGGCTTCCCAATGAATACTGGGGCTGGGGTGG 1104
QY 284 YGluAspAspAspLeuAlaThrArGTAr 293
DB 1105 TGAGGATGACGACATTTGCTACCAAGGTCA 1132

RESULT 6
US-10-750-623-30398
; Sequence 30398, Application US/10750623
; Publication No. US20050287531A1

```


GENERAL INFORMATION:
APPLICANT: MMI GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM1100-1
CURRENT APPLICATION NUMBER: US/10/750,623
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
NUMBER OF SEQ ID NOS: 64922
SOFTWARE: PatentIn version 3.1
SEQ ID NO 30398
LENGTH: 2507
TYPE: DNA
ORGANISM: Bovine 19866881075752
US-10-750-623-30398

Alignment Scores:
Pred. No.: 5,316-30 Length: 2507
Score: 331.50 Matches: 72
Percent Similarity: 50.53% Conservative: 24
Best Local Similarity: 37.89% Mismatches: 50
Query Match: 16.32% Indels: 44
Gaps: 3
DB: 6

US-10-661-430-1 (1-383) x US-10-750-623-30398 (1-2507)

QY 12 LysAspCysValAlaArgHisArgValAlaIleIleValProTyrArgAspArgGluAla 161
DB 563 GAAATGTCCTTCTCTTATCTTATCCCATPACTCTTGACTTCCCTCCATTC 642
QY 162 HisLeuArgGluMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAspTyr 181
DB 643 ACTTTATGCTCTCTCTCCAC-----GCCACCTATAGACTGATG 684
QY 182 AlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsn 201
DB 685 TCCTCTCCATATATCTTCCAGGCTGGAAATGGAACATTAAACAGGGAACCTGCTGAT 744
QY 202 Val-GlyTyrAspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspVal 221
DB 745 GTTCGGGGTGGCCGAGGCCCTGCGAGATGAGATGGAGCTGTTGTTCTGTCATGATGT 804
QY 221 AspLeuLeuProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHis 240
DB 805 GGATCTCTCGCCAGAGAACACACACCTGATGTGTGTCGATCCCGAGGACCTCGGCA 864
QY 240 sMetSerValAlaIleAspLysPheAsnTyrLys----- 251
DB 865 CGTTCCTGTTGCCATGAAACAAATTGGATACAGTAGAGGCGACAGAGGTAATCGAAG 924
QY 251 ----- 251
DB 925 CAGGGTATTCACCATTCAGTGGAAAGAGGAGATGCTCCCAAAAGCCTGTGGATTC 984
QY 252 -----LeuProTyrSerAlaIlePheGlyGlyIleSerAlaIle 264
DB 985 CAAGATCTTCTTCCCGAGCTCCCGATATCCCAAGTACTTTGGCGGGGTCTCAGCGGT 1044
QY 264 uThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTyrGlyIle 284
DB 1045 CACTCTCTGACAGTACCTGAAAGATGATGCTTCCCAATGATATCTGGGCGTGGGGTGG 1104
QY 284 YGluAspAspAspLeuAlaThrArgThr 293
DB 1105 TGAGGATGACGACATTGCTACCAAGGTCA 1132
RESULT 7

US-10-995-561-13225/c
Sequence 13225, Application US/10995561
Publication No. US20050272054A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ ID NOS: 85702
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13225
LENGTH: 16822
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(16822)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-;
US-10-995-561-13225

Alignment Scores:
Pred. No.: 2,216-26 Length: 16822
Score: 313.00 Matches: 68
Percent Similarity: 50.00% Conservative: 13
Best Local Similarity: 41.98% Mismatches: 36
Query Match: 15.41% Indels: 46
Gaps: 2
DB: 6

US-10-661-430-1 (1-383) x US-10-995-561-13225 (1-16822)

QY 185 lIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyr 204
DB 13662 ATTATCTTCCAGGCTGGAAGAAATGGAACATTAAACAGGGAACCTGTTGAAAGTTGGGGTG 13603
QY 205 AspValAlaSerArgLeuTyrProTyrGlnCysPheIlePheHisAspValAspLeuLeu 224
DB 13602 CGAGAGGCCCTCGCGATGAAAGATGGGACTGCTGTTCTTGACAGATGTGACCTCTTG 13543
QY 225 ProGluAspAspArgAsnLeuTyrThrCys---ProIleGlnProArgHisMetSerVal 243
DB 13542 CCAGAAATAGACCAACATCTGATGTGTGACCCCGGGAGCCCGGCATGTCGGT 13483
QY 244 AlaIleAspLysPheAsnTyrLys----- 251
DB 13482 GCTATGAACAAGTTTGTGATACAG-GTAGAGGCGATGGAAGGTAATCTGGGAAAACAGRAAG 13424
QY 251 ----- 251
DB 13423 CCTACTATCTGTGGAAGAGAGAGCCAGCGATTTGTGGGATGTGTCTCCCGAGA 13364
QY 252 -----LeuProTyrSerAlaIlePheGly 259
DB 13363 AGACTCGGGGTATGCAAGATCTTCTCTCCCTAGCTCCGTAACCCCAAGTACTCGGA 13304
QY 260 GlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPhe 279
DB 13303 GGAGTCTGACGACATTAATCTCCGACCACTGAAAGATGCTTCCCAATGAAATAC 13244
QY 280 TrrGlyTrrGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLys 299
DB 13243 TGGGGCTGGGGTGGTGAAGATGACGACATTTGCTACAGGTGACAGCTTCTGCACTTCT 13184
QY 300 ValSer 301
DB 13183 CCCTCC 13178
RESULT 8
US-10-995-561-13401
Sequence 13401, Application US/10995561
Publication No. US20050272054A1


```
/ CURRENT APPLICATION NUMBER: US/10/750,185
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 34973
/ LENGTH: 3458
/ TYPE: DNA
/ ORGANISM: Bovine 19866881077545
US-10-750-185-34973

Alignment Scores:
Pred. No.: 1.87e-05 Length: 3458
Score: 136.50 Matches: 30
Percent Similarity: 64.91% Conservative: 7
Best Local Similarity: 52.63% Mismatches: 19
Query Match: 6.72% Indels: 2
DB: Gaps: 1

US-10-661-430-1 (1-383) x US-10-750-185-34973 (1-3458)

QY 188 GlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyr--AspVal 206
Db 3288 CAGCAGCGGTAGAGACACCTTCAACCGGCCCAAGCTGCTCAACGGGTTCCTAGAGGCA 3347
QY 207 AlaSerArgLeuTyrProTdpGlnCysPheIlePheHisAspValAspLeuLeuProGlu 226
Db 3348 CTCAAGAGAGACTCCACCTTCAACTGCTTCACTTCACTGTAAGTGAAGTGAAGTCCCATG 3407
QY 227 AspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSerVal 243
Db 3408 GATGACCGCAACTGTCACGCTGTGG-TACACAGCCCGCCACCTTGCCATT 3457

RESULT 14
US-10-750-623-34973
/ Sequence 34973, Application US/10750623
/ Publication No. US20050287531A1
/ GENERAL INFORMATION:
/ APPLICANT: MMI GENOMICS, INC.
/ APPLICANT: DENISE, Sue K.
/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen
/ APPLICANT: PANTIN, Dennis
/ TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
/ FILE REFERENCE: MM11100-1
/ CURRENT APPLICATION NUMBER: US/10/750,623
/ CURRENT FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437,482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 34973
/ LENGTH: 3458
/ TYPE: DNA
/ ORGANISM: Bovine 19866881077545
US-10-750-623-34973

Alignment Scores:
Pred. No.: 1.87e-05 Length: 3458
Score: 136.50 Matches: 30
Percent Similarity: 64.91% Conservative: 7
Best Local Similarity: 52.63% Mismatches: 19
Query Match: 6.72% Indels: 2
DB: Gaps: 1

US-10-661-430-1 (1-383) x US-10-750-623-34973 (1-3458)

QY 188 GlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnValGlyTyr--AspVal 206
Db 3288 CAGCAGCGGTAGAGACACCTTCAACCGGCCCAAGCTGCTCAACGGGTTCCTAGAGGCA 3347
```

```
QY 207 AlaSerArgLeuTyrProTdpGlnCysPheIlePheHisAspValAspLeuLeuProGlu 226
Db 3348 CTCAAGAGAGACTCCACCTTCAACTGCTTCACTTCACTGTAAGTGAAGTGAAGTCCCATG 3407
QY 227 AspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSerVal 243
Db 3408 GATGACCGCAACTGTCACGCTGTGG-TACACAGCCCGCCACCTTGCCATT 3457

RESULT 15
US-11-000-688-1051
/ Sequence 1051, Application US/11000688
/ Publication No. US20050287544A1
/ GENERAL INFORMATION:
/ APPLICANT: BERTUCCI, Francois
/ APPLICANT: HOUIGATTE, Remi
/ APPLICANT: BIRNBAUM, Daniel
/ TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
/ FILE REFERENCE: 1423-R-03
/ CURRENT APPLICATION NUMBER: US/11/000,688
/ CURRENT FILING DATE: 2004-12-01
/ PRIOR APPLICATION NUMBER: US 60/525,987
/ PRIOR FILING DATE: 2003-12-01
/ NUMBER OF SEQ ID NOS: 1596
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1051
/ LENGTH: 3745
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial sequences: primer
/ NAME/KEY: misc feature
/ LOCATION: (1)-(3745)
/ OTHER INFORMATION: chondroitin sulfate galnact-2 (GALNACT-2)
US-11-000-688-1051

Alignment Scores:
Pred. No.: 0.0497 Length: 3745
Score: 109.50 Matches: 98
Percent Similarity: 34.57% Conservative: 61
Best Local Similarity: 21.30% Mismatches: 163
Query Match: 5.39% Indels: 138
DB: Gaps: 20

US-10-661-430-1 (1-383) x US-11-000-688-1051 (1-3745)

QY 30 LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIleAlaAspVal 49
Db 747 AAACATCCCGAGTATGGGGTCAATTCCTTGAAGAATTTCCTTA-----ATG 797
QY 50 AspAlaMetGluAlaValLeuGlyAsnThr-----Ala 60
Db 798 AAAGATTTCATATGGAATGGAATGCTCAGCTCCGCAATCTGAAGAAAGCCAGTTAGAAA 857
QY 61 SerThrSerAspAspLeuAspThrTrpAsnSerThrPheSerProIleSerGluVal 80
Db 858 GACAAAGAGATGAATTTGTCGGAAGTTTGAAGCGGGCTGAGAGTCATTAATTCCT 917
QY 81 AsnGlnThrSerPheMetGluAspIle-----ArgProIleLeuPhePro 95
Db 918 GATGAAGATGATGAACAGAAAGATGAGAGGCTCCCTTGAAGAAAGATGATTTTAAT 977
QY 96 AspAsnGlnThrLeuGlnPheCysAsnGlnThrProProHisLeuValGlyProIleArg 115
Db 978 GAAAGTACTTTCGTAGAAGTTATTTATCGCACTGAGAGAGATGAAGGACACAGTAA 1037
QY 116 ValPheLeuAspGluProAsp-----Phe 123
Db 1038 CTCCTTTTAAAGAAAGACCTTACGAAATATAGACATGTGACCTCTTCGCGCTTTT 1097
QY 124 LysThrLeuGlnLysIleTyrProAspThrHisAlaGlyHisIleValMetProLysAsp 143
```

Db	1098	GGACCTCTCAGAAAGTGAAGAGTCAGATGATTCAGC-----	1133
Qy	144	CysValAlaArgHisArgValAlaIleIleValProTyrArgAspArg--GluAlaHis	162
Db	1134	---ATCATGATCAATATATATTAATATCATTTGTGCCACTTGGCAAGAACTGAAGCATTTT	1190
Qy	163	LeuArgIleMet-----LeuHisAsn	169
Db	1191	GTACAAATTTATGACAGAACTTCAGGAGTGTGTATTCATCAAGACAAAGAAATTCATCTC	1250
Qy	170	LeuHisSerLeuLeuAlaIysGlnGlnIleAspTyrAlaIlePheIleValGlnIVal	189
Db	1251	ACAGTGGTGATTTTGTGTAAGAAGAGACTGTCTAAAGCAAGTCTATCTTGAAATCTGTCTC	1310
Qy	190	Ala-----AsnGlnThrPheAsnArg	196
Db	1311	ACCAATGAGTCAATATTTTCACAATTTACACTTGGTCTCATTTGAATGAAGAAATTTAATGCT	1370
Qy	197	GlyIysIleuMetAsnValGly-----TyrAspValAlaSerArgLeuTyrProTrp	213
Db	1371	GGACGAGGACTCAATATGCTGGTGGTCCCGACCTTGGCAAGAGAGAGCTCTTG-----	1421
Qy	214	GlnCysPheIlePheHisAspValAsp-----	222
Db	1422	-----ANGTTTTTCTGATGCTTGATATCTATTTCTCAGCCGAATTCCTTAACAGCTGC	1475
Qy	223	---LeuLeuProGluAspArgArgAsnLeuTyrThrCysProIleGlnProArgHisMet	241
Db	1476	CGGTTAAATGCTGACGACGATGACAGAGGCTTTT---TACCCTGGGAGTTCAGCTTTTAC	1532
Qy	242	SerValAlaIleAspIysPheAsnTyrIleLeuPro-----	253
Db	1533	AATCTGCGCATTTGTTATGCCAAACAGGAAGGCCAACCACTGTGGAGCAGCAGCTGCTT	1592
Qy	254	-----TyrSerAlaIlePheGly---GlyIleSerAlaLeuThrIys	266
Db	1593	CACAAAAGAGATTCGCTTTTGGCGAATTTTGGCTTTGGAAAGACTGTTCAGATGCT	1652
Qy	267	AspHisIleuIysIysIleAsnGlyPheSerAsnAspPheTrpGlyTrpGlyGlyAsp	286
Db	1653	TCAGATTTCTGACCACTTGCTGGATTGTGACATGGAAGTGAAGAAAGTTGGGGGTGGAGAAGAT	1712
Qy	287	-----AspAspLeuAlaThrArgThrSerMetAla	296
Db	1713	GTTCATCTTTATCCAAAATTACTTACATGCTGATCTCATTTGATTCGAGCTCGGGTTCCT	1772
Qy	297	GlyIleuIysValSerArgTyrProThrGlnIleAla-----ArgTyr	310
Db	1773	GGTCTTTTCCACCTCGCATGGAAGAGCGTGTGTGATGAGCTGACCCCGAGACAGATAC	1832
Qy	311	LysMetIle-----LysHisSerThrGluAlaThrAsn-ProValAsnIysCys--	326
Db	1833	CGCATGTGCATCAAGTCTAAAGCCATGATGAGGCTCTCCATCCCACTGGAGATGCTG	1892
Qy	327	-----ArgTyrIysIleMetGlyGlnThr	334
Db	1893	GTCTTCCAGGAGAAATGAGACGCACTTTCAATAACAGGCAT-----ACAGACAAAC	1946
Qy	334	TlysArgArgTrpThrArgAspGlyLeuSerAsnIleuIysTyrIlysIleuValAsn----	352
Db	1947	AGTGAAGCTGTGGTGAATCATCATTAATTAAGCGTTACTGTATGTAACCAACAAACAGACAC	2006
Qy	353	-LeuGlnIleuIysProLeuTyrThrArgAlaValValAspIleuIleuGlyIysAspCys	371
Db	2007	TATTTATTTAAGCTTACTTCTTCACTTCCAGATGCACTGCTCTTTTGGAGAAAGACATGT	2064

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:23:20 ; Search time 28 Seconds
(without alignments)
116.088 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVRLKSLVLCNV.....VDLKKCRRLRDPPTCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 61072 seqs, 8486849 residues

Total number of hits satisfying chosen parameters: 61072

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_New:*
1: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep:*
5: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	642.5	31.6	344	US-11-102-240-42	Sequence 42, Appl
2	100.5	4.9	532	US-11-102-240-72	Sequence 72, Appl
3	87	4.3	885	US-10-055-877-240	Sequence 240, Appl
4	87	4.3	885	US-10-055-877-241	Sequence 241, Appl
5	83	4.1	506	US-11-134-241-31	Sequence 31, Appl
6	83	4.1	603	US-11-134-241-43	Sequence 43, Appl
7	82	4.0	403	US-10-821-234-1490	Sequence 1490, Appl
8	80.5	4.0	552	US-10-131-826A-196	Sequence 196, Appl
9	80.5	4.0	2080	US-10-821-234-1640	Sequence 1640, Appl
10	80	3.9	269	US-11-134-241-27	Sequence 27, Appl
11	80	3.9	366	US-11-134-241-39	Sequence 39, Appl
12	78	3.8	532	US-10-793-626-236	Sequence 236, Appl
13	77.5	3.8	704	US-11-172-145-25	Sequence 25, Appl
14	77.5	3.8	965	US-11-172-145-4	Sequence 4, Appl
15	77.5	3.8	922	US-11-124-215-3	Sequence 3, Appl
16	77	3.8	458	US-11-016-503-16	Sequence 16, Appl
17	77	3.8	458	US-11-089-803-6	Sequence 6, Appl
18	77	3.8	458	US-11-149-738-2	Sequence 2, Appl
19	77	3.8	458	US-11-155-269-2	Sequence 2, Appl
20	77	3.8	2204	US-11-052-554A-174	Sequence 174, Appl
21	77	3.8	2335	US-10-821-234-1610	Sequence 1610, Appl
22	76.5	3.8	385	US-10-392-234A-28	Sequence 28, Appl
23	75	3.7	858	US-10-645-441-18	Sequence 18, Appl
24	75	3.7	858	US-10-645-441-20	Sequence 20, Appl
25	75	3.7	858	US-10-645-441-23	Sequence 23, Appl

26	75	3.7	896	US-10-467-657-7004	Sequence 7004, Appl
27	74.5	3.7	458	US-11-016-503-12	Sequence 12, Appl
28	74.5	3.7	458	US-11-089-803-2	Sequence 2, Appl
29	74.5	3.7	513	US-10-878-556A-112	Sequence 112, Appl
30	74	3.6	561	US-10-454-437-330	Sequence 330, Appl
31	72.5	3.6	430	US-11-016-503-17	Sequence 17, Appl
32	72.5	3.6	654	US-10-858-730-5	Sequence 5, Appl
33	72.5	3.6	621	US-10-632-150-28	Sequence 28, Appl
34	72.5	3.6	621	US-11-073-457-28	Sequence 28, Appl
35	72.5	3.6	621	US-11-073-460-28	Sequence 28, Appl
36	72.5	3.6	4868	US-11-044-111-24	Sequence 24, Appl
37	72	3.5	535	US-11-134-241-29	Sequence 29, Appl
38	72	3.5	631	US-11-134-241-41	Sequence 41, Appl
39	72	3.5	914	US-10-312-954-2	Sequence 2, Appl
40	72	3.5	1255	US-11-052-554A-273	Sequence 273, Appl
41	71.5	3.5	774	US-11-070-627-7	Sequence 7, Appl
42	71.5	3.5	1062	US-10-821-234-1079	Sequence 1079, Appl
43	71	3.5	376	US-10-454-437-412	Sequence 412, Appl
44	71	3.5	775	US-10-131-826A-120	Sequence 120, Appl
45	70.5	3.5	415	US-11-110-851-64	Sequence 64, Appl

ALIGNMENTS

```
RESULT 1
US-11-102-240-42
; Sequence 42, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 42
; LENGTH: 344
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-42
Query Match          31.6%  Score 642.5; DB 7; Length 344;
Best Local Similarity 38.5%  Pred. No. 3.9e-55;
Matches 143; Conservative 55; Mismatches 124; Indels 49; Gaps 8;

QY      5 HLA-A-RKSLVLCVLL-----LVHAMIYKIPISLYENLTIGSSTLIADVDAMEAVL 56
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      8 HLYSFRLLLLTLLTCLTYVGATSNVYFGA-IQELPKAKFEFANPHKTLIL---GKGKTL 63
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      57 GNTASTSDLLDPTMNSTSPISSEVNGTSPFMDIRLPDPNQTLOPCNQTPHLVGPPIRV 116
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
DB      64 TNEASTKKEVELDN-----CEVSFYLKGQSKL 90
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      117 FLDEDPFTLEKTYDTH--AGHGMPKDCVARRVATIVPYRDEARLRTMLHNLHSL 174
      :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
DB      91 IF-KPDL-TLSEVQAKNKVSRGRYRPOECKALQVALIVPHRNRKILMTLEHLPPL 148
      :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||

QY      175 AKQOLDVAIFIVEQVANQTFNRGKLMNVGYDVASRLYPMQCFIFHDVDLLEDDENLYTC 234
      :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||  :||
```

Dd	149	ORQOUDYGIYVHNOABGKKFNRAKLLNVGYLBAKKEBNMDCFFHNDVDLVBPENDFNLKYC	208
Qy	235	PIQPRHMSVAIDKFNYKFLPYSAIFGGSIALTKDHLKKNFSDNFWMGEGEDDLATRTS	294
Dd	209	EEHPHNLVVGGRNSTGRLYKSGYFGGVTMLSRQFPKNGFSNNYMGWGEDDDLRLRAVE	268
Qy	295	MAGLVSRYPTQIARYKMIKHSTEATNPVNKCXYKTMGQTRKWTBDGLSNLKYKLVNLB	354
Dd	269	LORMKISRPLPEVGKTYWTFHTEDDKGNBEVAEMKLLHQSRYWRDTGLSSCSQKLVSE	328
Qy	355	LKELYTRAVVD	365
Dd	329	NNPLYNITVD	339

```

RESULT 2
US-11-102-240-72
; Sequence 72, Application US/11102240
; Publication No. US20050260647A1
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIPODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESS
; TITLE OF INVENTION: SEOPHAGEAL TUMOR
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 72
; LENGTH: 532
; TYPE: PRT
; ORGANISM: Homo Sapien
US-11-102-240-72

```

Query Match	Similarity	4.9%	Score 100.5	DB 7	length 532
Best Local	Similarity	22.1%	Pred. No. 0.043		
Matches	Conservative	34	Mismatches	93	Indels 67; Gaps 11
Qy	145	VARRHVAIVPYDRBAHLRIMLNLHSL	-----	LAKQOLDVAIFTEQVA	190
		: : : : :		: : : : :	
Db	258	MANTLINIVPLAKRVDRKFRQMNFRREMCIEBODGCVHLTVVYFGKEELINEVKGILLENTS	317		
Qy	191	-----NCFNFGKLMNNGYDVAASLYPMQ-----CFIFHDVLLLEEDDRNLYT	233		
		: : : : :		: : : : :	
Db	318	KANERNFTFTIOLNGEFSRKG-----GLDVGARF-WKGSNVLILFFCDVDIYTTSR-PLNT	370		
Qy	234	C-----PIOPRHMSVAIDKENYKLPY---SAI-----FG-GISALTYK	266		
		: : : : :		: : : : :	
Db	371	CLNTQPGKKVFYFVLFSSQNGEIIYGHDDADVPLEQOLVIKKENGFWHDFSGMTCOYR	430		
Qy	267	DHLAKKINGNSNDFWCGGDDDLATPTSMAGKVSRSYPQIARLYKMIKHSSTATPYNNCK	326		
		: : : : :		: : : : :	
Db	431	SPFINITGGDLDDKGGMGEDVHLRYKYLHSLNLIIVATPVRGLFH-LMHKKSCMDBLTPE	488		
Qy	327	RYKIMGQTK	335		
		: : : : :		: : : : :	
Db	489	OYKCMQOSK	497		

RESULT 3
US-10-055-877-240
; Sequence 240, Application US/10055877

```

Publication No. US20050288241A1
GENERAL INFORMATION:
APPLICANT: Decristofaro, Marc
APPLICANT: Padigaru, Muralidhara
APPLICANT: Miller, Charles
APPLICANT: Tchernev, Velizar
APPLICANT: Zhong, Mei
APPLICANT: Anderson, David
APPLICANT: Ballinger, Robert
APPLICANT: Gerlach, Valerie
APPLICANT: Spytek, Kimberly
APPLICANT: Ratelli, Luca
APPLICANT: Kekuda, Ramesh
APPLICANT: Guo, Xiaojia
APPLICANT: Zernhusen, Bryan
APPLICANT: Andrew, David
APPLICANT: Mezes, Peter
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine
APPLICANT: Eisen, Andrew
APPLICANT: Wolenc, Adam
APPLICANT: Baumgartner, Jason
APPLICANT: Shimkets, Richard
APPLICANT: Gusev, Vladimir
APPLICANT: Verne, Corine
APPLICANT: Taupier Jr., Raymond
APPLICANT: Pena, Carol
APPLICANT: Shenoy, Suresh
APPLICANT: Li, Li
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Perence
TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
FILE REFERENCE: 21402-251
CURRENT APPLICATION NUMBER: US/10/055,877
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: 60/262,892
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: 60/263,598
PRIOR FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 60/263,799
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/264,117
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,139
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: 60/264,478
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 60/263,351
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/272,870
PRIOR FILING DATE: 2001-03-02
PRIOR APPLICATION NUMBER: 60/275,990
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/275,927
PRIOR FILING DATE: 2001-03-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 512
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 240
LENGTH: 885
TYPE: PRT
ORGANISM: Homo sapiens
US-10-055-877-240

```

	Query Match	4.3%	Score 87;	DB 6;	Length 889;
	Best Local Similarity	18.9%	Pred No	1.8;	
	Matches	61;	Conservative	45;	Mismatches 82; Indels 134; Gaps 15
Qy	114 IRVFLDEPDKLEKTYIPDTHAGCHGMKDCVARRRVAIYPRDRRAHRLMLNHLSTL	173			
	: :				
Dd	69 VRIFLTEPLSGELCTLYOE-----PRE-----LPPEER-----AQNVTRL	103			
Qy	174 LAKQDLDAIFIVEVAQTENGRGLMNVGVDSARSLYPWQCFFIRHDVLLPEDDRNLVT	233			


```
Db 104 LKQ-----AEEVNT-----WEKLNHSADWQKIDETLE- 134
Qy 234 CPIOGRHNSVAIDKKNYLPYSAIFG-----ISALTQDHUKKINGFNDWGMCG 284
Db 135 ---RLQELQEADELDELKRLQAEVIGKSWQPVGDLLISL-ODHLEKVALR-----G 183
Qy 285 E-----DDDLATRTSMAGLKVSRYPQI-----ARYKMIGHSTE----- 318
Db 184 ELAPLKENVSHNDLRLQRLTIGIQLSPYNLSTEDLNTFWLQVAVEDRQLHEAHR 243
Qy 319 -----ATNPVNCKRYKIMGOTKRR-WTRDGLSNLKYLVNL----- 353
Db 244 DFGPASQHFLLSTSVQGPWERAISP-NKVPPYINHETQTCMDHPKMTLEYQSLADLNNVR 302
Qy 354 -----ELKPLYTRAVDIL 367
Db 303 FSAVRTAMKLRRLQKALCLDIL 324
```

RESULT 4

```
US-10-055-877-241
; Sequence 241, Application US/10055877
; Publication No. US20050288241A1
; GENERAL INFORMATION:
; APPLICANT: Decristofaro, Marc
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Miller, Charles
; APPLICANT: Tchernev, Vellizar
; APPLICANT: Zhong, Mei
; APPLICANT: Anderson, David
; APPLICANT: Ballinger, Robert
; APPLICANT: Gerlach, Valerie
; APPLICANT: Spyrek, Kimberly
; APPLICANT: Rattelli, Luca
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Guo, Xiaojia
; APPLICANT: Zernusen, Bryan
; APPLICANT: Andrew, David
; APPLICANT: Mezes, Peter
; APPLICANT: Patuturajan, Meera
; APPLICANT: Burgess, Catherine
; APPLICANT: Eissen, Andrew
; APPLICANT: Wolenc, Adam
; APPLICANT: Baumgartner, Jason
; APPLICANT: Shinkets, Richard
; APPLICANT: Gusev, Vladimir
; APPLICANT: Vernet, Corine
; APPLICANT: Taupier Jr., Raymond
; APPLICANT: Pena, Carol
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Casman, Stacie
; APPLICANT: Bolog, Ference
; TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoded Thereby
; FILE REFERENCE: 21402-251
; CURRENT APPLICATION NUMBER: US/10/055,877
; PRIOR APPLICATION NUMBER: 60/262,892
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/263,598
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: 60/263,799
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/264,117
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,139
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/264,478
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/263,351
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/272,870
```

```
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/275,990
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/275,927
; PRIOR FILING DATE: 2001-03-14
; Remaining Prior Application data removed - See file wrapper or PALM.
; NUMBER OF SEQ ID NOS: 512
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 241
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-055-877-241
```

```
Query Match 4.3% Score 87; DB 6; Length 885;
Beet Local Similarity 18.9%; Pred. No. 1.8;
Matches 61; Conservative 45; Mismatches 82; Indels 134; Gaps 15;
```

```
Qy 114 IRVFLDEPDPFTLEKIYPTDTHAGHGMPKDCVARRHVAIIVPYRDREAHRLIMLNHSL 173
Db 69 VRIFLEQPLBGLKLYQE-----PR-----LPPEER-----AQNVTRL 103
Qy 174 LAKQQLDYAIFIVEQVANQTFNRGKLMNVGYDVASRLYPWOCFTFHDVVDLPEDDRLNYT 233
Db 104 LKQ-----AEEVNT-----WEKLNHSADWQKIDETLE- 134
Qy 234 CPIOGRHNSVAIDKKNYLPYSAIFG-----ISALTQDHUKKINGFNDWGMCG 284
Db 135 ---RLQELQEADELDELKRLQAEVIGKSWQPVGDLLISL-ODHLEKVALR-----G 183
Qy 285 E-----DDDLATRTSMAGLKVSRYPQI-----ARYKMIGHSTE----- 318
Db 184 ELAPLKENVSHNDLRLQRLTIGIQLSPYNLSTEDLNTFWLQVAVEDRQLHEAHR 243
Qy 319 -----ATNPVNCKRYKIMGOTKRR-WTRDGLSNLKYLVNL----- 353
Db 244 DFGPASQHFLLSTSVQGPWERAISP-NKVPPYINHETQTCMDHPKMTLEYQSLADLNNVR 302
Qy 354 -----ELKPLYTRAVDIL 367
Db 303 FSAVRTAMKLRRLQKALCLDIL 324
```

RESULT 5

```
US-11-134-241-31
; Sequence 31, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Hu, Yi
; APPLICANT: Kleke, James Alvin
; APPLICANT: Potter, David George
; TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
; FILE REFERENCE: LEX-0144-USA
; CURRENT APPLICATION NUMBER: US/11/134,241
; PRIOR APPLICATION NUMBER: US/10/364,774
; PRIOR FILING DATE: 2003-02-11
; PRIOR APPLICATION NUMBER: US/09/795,926
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,920
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 60/186,558
; PRIOR FILING DATE: 2000-03-02
```

PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 31
LENGTH: 506
TYPE: PRT
ORGANISM: homo sapiens
US-11-134-241-31

Query Match 4.1%; Score 83; DB 7; Length 506;
Best Local Similarity 17.5%; Pred. No. 2;
Matches 49; Conservative 41; Mismatches 80; Indels 110; Gaps 11;

QY 72 STFSPISEVNQTSFMDIRPIILFPDNOQLQFCNQTPPHLVGPIRVFLDEPDEFTLEK--- 128
DB 51 STIIPFHNGWSSLLRTVHSV-----NRSPELVAEIVLVDDPSDREHLKKPLR 100
QY 129 ----IYP-----DTH--AGGHGMPK--DCVAH 148
DB 101 DYMAIFPSVRIILRTKKREGILRTRMIGASVATGDIVITFLDSHCENAVNWLPELLDRIAN 160
QY 149 RVAIIVPYRDREAHRLIMLHNSLLAKQQLDYAIFIVEQVANOTFNRGKLMNVGYDVAS 208
DB 161 RKTIVCEPID-----VIDHDFRYETQA-----GDAMRGAFD--- 192
QY 209 RLYPMQCFIFHDVDLLEP-----DDRNLVYTCPIQPRHMSVAIDKFNYKLPYSALFGGISAL 264
DB 193 ----WEMY-YKRIPIPELQKADPSDPFESPV-----MAGGLFAV 227
QY 265 TKDHLKKNFNSNDFMGCGEDDLATRTSMAGLKVSRY 304
DB 228 DRKFWELGIGYDPLGLEIWGGEQYEISFKVMCGGRMEDIP 267

RESULT 6

US-11-134-241-43
Sequence 43; Application US/11134241
Publication No. US20050287566A1
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Halbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Wilganowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kieke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSERASE PROTEINS AND
FILE REFERENCE: LEX-0144-USA
CURRENT APPLICATION NUMBER: US/11/134,241
CURRENT FILING DATE: 2005-05-20
PRIOR APPLICATION NUMBER: US/10/364,774
PRIOR FILING DATE: 2003-02-11
PRIOR APPLICATION NUMBER: US/09/795,926
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 43
LENGTH: 603
TYPE: PRT
ORGANISM: homo sapiens

US-11-134-241-43

Query Match 4.1%; Score 83; DB 7; Length 603;
Best Local Similarity 17.5%; Pred. No. 2,6;
Matches 49; Conservative 41; Mismatches 80; Indels 110; Gaps 11;

QY 72 STFSPISEVNQTSFMDIRPIILFPDNOQLQFCNQTPPHLVGPIRVFLDEPDEFTLEK--- 128
DB 148 STIIPFHNGWSSLLRTVHSV-----NRSPELVAEIVLVDDPSDREHLKKPLR 197
QY 129 ----IYP-----DTH--AGGHGMPK--DCVAH 148
DB 198 DYMAIFPSVRIILRTKKREGILRTRMIGASVATGDIVITFLDSHCENAVNWLPELLDRIAN 257
QY 149 RVAIIVPYRDREAHRLIMLHNSLLAKQQLDYAIFIVEQVANOTFNRGKLMNVGYDVAS 208
DB 258 RKTIVCEPID-----VIDHDFRYETQA-----GDAMRGAFD--- 289
QY 209 RLYPMQCFIFHDVDLLEP-----DDRNLVYTCPIQPRHMSVAIDKFNYKLPYSALFGGISAL 264
DB 290 ----WEMY-YKRIPIPELQKADPSDPFESPV-----MAGGLFAV 324
QY 265 TKDHLKKNFNSNDFMGCGEDDLATRTSMAGLKVSRY 304
DB 325 DRKFWELGIGYDPLGLEIWGGEQYEISFKVMCGGRMEDIP 364

RESULT 7

US-10-821-234-1490
Sequence 1490; Application US/10821234
Publication No. US20050255144A1
GENERAL INFORMATION:
APPLICANT: Labat, Ivan
APPLICANT: Stache-Crain, Birgit
APPLICANT: Andarmani, Susan
APPLICANT: Tang, Y. Tom
TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
FILE REFERENCE: 821A
CURRENT APPLICATION NUMBER: US/10/821,234
CURRENT FILING DATE: 2004-04-07
PRIOR APPLICATION NUMBER: US 60/462,047
PRIOR FILING DATE: 2003-04-07
NUMBER OF SEQ ID NOS: 1704
SOFTWARE: PC SEQ_genes Version 1.0
SEQ ID NO 1490
LENGTH: 403
TYPE: PRT
ORGANISM: Homo sapiens
US-10-821-234-1490

Query Match 4.0%; Score 82; DB 6; Length 403;
Best Local Similarity 21.5%; Pred. No. 1,8; Indels 116; Gaps 19;
Matches 76; Conservative 45; Mismatches 117;

QY 76 PISEVNQTSFMDIRPIILFPDNOQLQFCNQTPPHLVGPIRVFLDEP---DEFTL----- 126
DB 63 PGSKVKKKEVFEAVTIV-----ETPRMVVVGIVGETRGARTFETVPAEHI 110
QY 127 ----EKIYDPTIAG-----HGPKKC-----VABHVAII 153
DB 111 SDECKRRFYKNMHSKKKAFTKCKKQODEGKQLEKOFSSMKKCYQVIRIAHTQMRL 170
QY 154 VPYRDREAHRLIMLHNSLLAKQQLDYAIFIVEQ--VANQTFNRGLMNV-----GY 204
DB 171 LPIRQKKAHL--MEIQVNGTVAEKIDMARERLEQVPPNVQVQODEMIDVIGTGTGKG 228
QY 205 -DVASRLYPWQCFIFHDVDLLEPDDRNL--YTC-----PIQPR-----HMSVAI 245
DB 229 KGYTSR-----WHTKKLPRTKTRGLRKVACIGAMHPARVAFVSARAGQKGYHRTET 280
QY 246 DKFNKLPYSALIFGGSALTCKDHLKKNFNSNDFMGCGEDDLATRT--SMAGLKVSRY 303
DB 281 NKKIYKI-----GGGYLIKDGKLIKNNASTDY-----DLSDKSINPLGSP--VHY 323

QY 304 PTOIARYKMKIKHSTATNP-VNCKRYKIMGOTRKRWRTRDGLSNLKYLVNLEIK 356
Db 324 GEVYTDPMWKKCGVGTKKRVLTLRKSLVOTRR---ALEKIDUKFIDTTSK 373

RESULT 8

US-10-131-826A-196
; Sequence 196, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Flvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerlitsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131, 826A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 196
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-131-826A-196

Query Match 4.0%; Score 80.5; DB 6; Length 552;
Best Local Similarity 22.0%; Pred. No. 4;

Matches 54; Conservative 36; Mismatches 77; Indels 79; Gaps 14;

QY 82 QTSFMEIDPILFLPDNQTLOFCNQTPPHLVGPPIV---FLDEP-DPTELEKIYPTDHAGG 137
Db 124 RSTLARTISVL-----NRTPTHLIREIIVDDSDNDPDDCKQILK----- 164
QY 138 HGMFK-DCVARRHVAIIVPYRD-----EAHIRMVLAHNLHSLAKQQLDYAI 183
Db 165 --LPKVKCIARNRERQGLVRSIRIGADIAGTTLTFLDSHCEVNRDWLOPLHVRKEDYTR 222

QY 184 F---IVEQVANOTFN-----RGKLMNVGYDVASRLYPWQCFIFHDVLLPEDD-RNL 231
Db 223 YVCEPIDIINDLTFYISASRLG-----GFD-WSLRFQW-----QLSPEQKARL 269
QY 232 YTCPIQPRHMSVAIDKFNKYLPYSAIFGISALTYDHLKTKNGFSDNFMWCGEDDDLAT 291
Db 270 --DPTEPIRTPI-----IAGGLFVIDKAMFEDYLGKYMMDMDIMWGENFEISF 314

QY 292 RTSMAG 297
Db 315 RVWMCg 320

RESULT 9

US-10-821-234-1640
; Sequence 1640, Application US/10821234
; Publication No. US2005025114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andermani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821, 234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462, 047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: pc_seq_genes Version 1.0
; SEQ ID NO 1640
; LENGTH: 2080
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1640

Query Match 4.0%; Score 80.5; DB 6; Length 2080;
Best Local Similarity 21.7%; Pred. No. 26;
Matches 53; Conservative 31; Mismatches 73; Indels 87; Gaps 12;

QY 38 LITGSSTLIADVDAMEAVLGNTASTSDLLDTWNSPFSISVNOT-SFME-DIRPIFLP 95
Db 1450 LSPGEDVIL-DIDDEPLI---PIQEEFIDWMSKFASIGREKCGSYLEKDFTLKYV 1505
QY 96 DNG-----TLQFCN-----QTPPHLVGIRFLDEPDKTELEKIYPTH 134
Db 1506 DTQLNVEAFEGLSDFCNTFKLYRGKTOETEDPSVIG-----EFKGLFKIY--- 1553
QY 135 AGGHGMPKDCVARRHVAIIVPYRD-----REAHLRIMVLAHNLHSLAK----- 176
Db 1554 -----LPED-----PALPMPRQFHQLAAGQFQECVRIIYVRAFGLOPKDPNGKCDPY 1602
QY 177 -----QQLDYAIFIVEQVANOTFNRRGKLMNVGYDVASRLYPWQCFIFHDVLLP 225
Db 1603 IKISIGKKSVDQDNVYIPCTLEPVPFGKMFELCTLPLEKDLKITLY-----DYDLLS 1654
QY 226 EDDR 229
Db 1655 KDEK 1658

RESULT 10

US-11-134-241-27
; Sequence 27, Application US/11134241
; Publication No. US20050287568A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Hilbun, Erin
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Abuin, Alejandro
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.

APPLICANT: Walke, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kleke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
FILE REFERENCE: LEX-0144-USA
CURRENT APPLICATION NUMBER: US/11/134,241
CURRENT FILING DATE: 2005-05-20
PRIOR APPLICATION NUMBER: US/10/364,774
PRIOR FILING DATE: 2003-02-11
PRIOR APPLICATION NUMBER: US/09/795,926
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 269
TYPE: PRT
ORGANISM: homo sapiens
US-11-134-241-27

Query Match 3.9%; Score 80; DB 7; Length 269;
Best Local Similarity 17.9%; Pred. No. 1.6;
Matches 50; Conservative 41; Mismatches 75; Indels 114; Gaps 12;

QY 72 STFSPISEVNTQSFMEDIRILFPDNTQLQFCNQTPHVLGPIRVFLDEPDFLTLEK--- 128
DB 51 SIIFPHNEGWSLRTVHSV-----NRSPPELVAEIVLVDDFSDRHLKPLE 100
QY 129 ----IYP-----DTH--AGGHGMPK--DCVARR 148
DB 101 DYVALPFSVRLIKTKREGILIRTRMLGASVATGDIITFLDSHCANVNMPLPLLDRIARN 160
QY 149 RVAIIVPYRDREAHRLIMLNHLSLAKQOLDVAIFIVEQVANQTFNRGKLMNVGYDVA 208
DB 161 RKTIVCPMD-----VIDHDDFRYETQA-----GDAMRGARD--- 192
QY 209 RLYPWCQFIHVDLLPE---DDRNLYTCPIQPRHMSVAIDKENYKLPYSALFGISAL 264
DB 193 ----WEMY-YKRIPIPELQKADSPDPESPV-----MAGGLFAV 227
QY 265 TKDHLKKINGFSNDFWCGGEDDDLATRTSMAGLKVSRRP 304
DB 228 DRKFWELGGYDGLIHWGEGYEISFKVS---QLSRRP 263

RESULT 11
US-11-134-241-39
Sequence 39, Application US/11/134241
Publication No. US20050287568A1
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Halbun, Erin
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Abuin, Alejandro
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
APPLICANT: Walke, D. Wade
APPLICANT: Wilgowski, Nathaniel L.
APPLICANT: Hu, Yi
APPLICANT: Kleke, James Alvin
APPLICANT: Potter, David George
TITLE OF INVENTION: NOVEL HUMAN TRANSFERASE PROTEINS AND
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING THE SAME
FILE REFERENCE: LEX-0144-USA

CURRENT APPLICATION NUMBER: US/11/134,241
CURRENT FILING DATE: 2005-05-20
PRIOR APPLICATION NUMBER: US/10/364,774
PRIOR FILING DATE: 2003-02-11
PRIOR APPLICATION NUMBER: US/09/795,926
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/185,920
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: US 60/186,558
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: US 60/191,849
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 39
LENGTH: 366
TYPE: PRT
ORGANISM: homo sapiens
US-11-134-241-39

Query Match 3.9%; Score 80; DB 7; Length 366;
Best Local Similarity 17.9%; Pred. No. 2.5;
Matches 50; Conservative 41; Mismatches 75; Indels 114; Gaps 12;

QY 72 STFSPISEVNTQSFMEDIRILFPDNTQLQFCNQTPHVLGPIRVFLDEPDFLTLEK--- 128
DB 148 SIIFPHNEGWSLRTVHSV-----NRSPPELVAEIVLVDDFSDRHLKPLE 197
QY 129 ----IYP-----DTH--AGGHGMPK--DCVARR 148
DB 198 DYVALPFSVRLIKTKREGILIRTRMLGASVATGDIITFLDSHCANVNMPLPLLDRIARN 257
QY 149 RVAIIVPYRDREAHRLIMLNHLSLAKQOLDVAIFIVEQVANQTFNRGKLMNVGYDVA 208
DB 258 RKTIVCPMD-----VIDHDDFRYETQA-----GDAMRGARD--- 289
QY 209 RLYPWCQFIHVDLLPE---DDRNLYTCPIQPRHMSVAIDKENYKLPYSALFGISAL 264
DB 290 ----WEMY-YKRIPIPELQKADSPDPESPV-----MAGGLFAV 324
QY 265 TKDHLKKINGFSNDFWCGGEDDDLATRTSMAGLKVSRRP 304
DB 325 DRKFWELGGYDGLIHWGEGYEISFKVS---QLSRRP 360

RESULT 12
US-10-793-626-236
Sequence 236, Application US/10/793626
Publication No. US20050255478A1
GENERAL INFORMATION:
APPLICANT: KIMMERLY, WILLIAM JOHN
TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
FILE REFERENCE: PU3480US
CURRENT APPLICATION NUMBER: US/10/793,626
CURRENT FILING DATE: 2004-03-04
PRIOR APPLICATION NUMBER: 60/164,258
PRIOR FILING DATE: 1999-11-09
NUMBER OF SEQ ID NOS: 4472
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 236
LENGTH: 258
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: amino acid sequence
US-10-793-626-236

Query Match 3.8%; Score 78; DB 6; Length 258;
Best Local Similarity 23.7%; Pred. No. 2.4;
Matches 42; Conservative 24; Mismatches 47; Indels 64; Gaps 8;
QY 28 IYKIPSLYENLTIGSST----LIADVAMEAVLGNTASTSDLLDTWNSTFSPISFVNQ 82

Db 44 I E K I Y N L A G N I G I G H E T N K V L D T I F D R D M E M I A G N --- H D E A L --- M S L V N G 90
Qy 83 T S F M D I P I L E P D N O T I O --- F C N O T P H L V G P I R --- V F --- 117
Db 91 T Y P P E D L G K G F E H H Q M I E G H L D E S Y D E I N O L P R Y I E T I G K K I L F H Y I E N D R K S A 150
Qy 118 - L D E E D F K T L E K I Y P D T H A G H G M P C V A R R H V A I I V P Y R D R E A H L R I M L N - L H 171
Db 151 P I D E O P F A P I T K --- D D E O A I S E L F K X E A D L I I F G N H L H 189

RESULT 13
US-11-172-145-25
; Sequence 25, Application US/11172145
; Publication No. US20050272696A1
; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; TITLE OF INVENTION: GLYCOSAMINOGLYCAN POLYMERS
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; PRIOR FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/491,362
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 25
; LENGTH: 704
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-25

Query Match 3.8%; Score 77.5; DB 7; Length 704;
Best Local Similarity 22.9%; Pred. No. 11;
Matches 30; Conservative 17; Mismatches 49; Indels 35; Gaps 5;

Db 171 H S L A K Q O L D V A I F I V E Q V A N O T F N R G K L M --- N G Y D V A S R L Y P M O C F I H D V D L P E 226
Qy 275 H N I T A E Q F L I N D P - Y L I E S L P E T A T I N N P S I T S K G N I S I D --- W R L E H F K K T D --- 322
Db 227 D D R N L Y T G C I O P R H M S V A I D K F N Y K L P Y S A I F G I S A L T K D L K K I N G S N D P M G M G E D 286
Qy 323 --- N U R L C D S P P R Y S C --- G N V A F S K E W L N K V G M F D E S F N H M G E D 363
Db 287 D D L A T R T S M A G 297
Qy 364 V E F G Y R L P A K G 374

RESULT 14
US-11-172-145-4
; Sequence 4, Application US/11172145
; Publication No. US20050272696A1

; GENERAL INFORMATION:
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: METHODS OF SELECTIVELY TREATING DISEASES WITH SPECIFIC
; TITLE OF INVENTION: GLYCOSAMINOGLYCAN POLYMERS
; FILE REFERENCE: 3554.104
; CURRENT APPLICATION NUMBER: US/11/172,145
; PRIOR FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/584,442
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 10/642,248
; PRIOR FILING DATE: 2003-08-15
; PRIOR APPLICATION NUMBER: 60/404,356
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/479,432
; PRIOR FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: 60/491,362
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: 10/195,908
; PRIOR FILING DATE: 2002-07-15
; PRIOR APPLICATION NUMBER: 09/437,277
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: 60/107,929
; PRIOR FILING DATE: 1998-11-11
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 965
; TYPE: PRT
; ORGANISM: Pasteurella multocida
US-11-172-145-4

Query Match 3.8%; Score 77.5; DB 7; Length 965;
Best Local Similarity 22.9%; Pred. No. 17;
Matches 30; Conservative 17; Mismatches 49; Indels 35; Gaps 5;

Db 171 H S L A K Q O L D V A I F I V E Q V A N O T F N R G K L M --- N G Y D V A S R L Y P M O C F I H D V D L P E 226
Qy 275 H N I T A E Q F L I N D P - Y L I E S L P E T A T I N N P S I T S K G N I S I D --- W R L E H F K K T D --- 322
Db 227 D D R N L Y T G C I O P R H M S V A I D K F N Y K L P Y S A I F G I S A L T K D L K K I N G S N D P M G M G E D 286
Qy 323 --- N U R L C D S P P R Y S C --- G N V A F S K E W L N K V G M F D E S F N H M G E D 363
Db 287 D D L A T R T S M A G 297
Qy 364 V E F G Y R L P A K G 374

RESULT 15
US-11-124-215-3
; Sequence 3, Application US/11124215
; Publication No. US2005026460A1
; GENERAL INFORMATION:
; APPLICANT: DEANGELIS, PAUL
; TITLE OF INVENTION: DNA ENCODING HYALURONAN SYNTHASE FROM PASTEURELLA MULTOCIDA AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 3554.101
; CURRENT APPLICATION NUMBER: US/11/124,215
; PRIOR FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: 10/217,613
; PRIOR FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/283,402
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/080,414
; PRIOR FILING DATE: 1998-04-02
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 28

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:22:30 ; Search time 114 Seconds
(without alignments)
1403.760 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSILVLCANV.....VDLEKDCRRRLRDPPTCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 1867569 seqs, 417829326 residues
Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA_Main:
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubppaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	383	4	US-10-661-430-1 Sequence 1, Appl1
2	718	35.4	404	3	US-09-925-301-1072 Sequence 1072, Ap
3	710	35.0	393	4	US-10-132-652-5 Sequence 5, Appl1
4	710	35.0	393	6	US-11-105-796-5 Sequence 5, Appl1
5	709	34.9	403	6	US-11-097-143-18555 Sequence 18555, A
6	680.5	33.5	362	4	US-10-459-311-7 Sequence 7, Appl1
7	673.5	33.2	362	4	US-10-132-652-6 Sequence 6, Appl1
8	673.5	33.2	362	6	US-11-105-796-6 Sequence 6, Appl1
9	667.5	32.9	352	4	US-10-087-192-2049 Sequence 2049, Ap
10	666	32.8	385	4	US-10-466-941-6 Sequence 6, Appl1
11	666	32.8	385	5	US-10-870-635-6 Sequence 6, Appl1
12	666	32.8	398	5	US-10-761-435A-16 Sequence 16, Appl1
13	663.5	32.7	371	5	US-10-761-435A-20 Sequence 20, Appl1
14	660.5	32.5	397	5	US-10-713-970-21 Sequence 21, Appl1
15	660	32.5	372	4	US-10-287-226-82 Sequence 82, Appl1
21	646	31.8	344	6	US-11-105-796-3 Sequence 3, Appl1
22	642.5	31.6	344	3	US-09-989-722-236 Sequence 236, App
23	642.5	31.6	344	3	US-09-989-723-236 Sequence 236, App
24	642.5	31.6	344	3	US-09-989-727-236 Sequence 236, App
25	642.5	31.6	344	3	US-09-989-727-236 Sequence 236, App
26	642.5	31.6	344	3	US-09-989-731-236 Sequence 236, App
27	642.5	31.6	344	3	US-09-989-732-236 Sequence 236, App

28	642.5	31.6	344	3	US-09-991-073-236 Sequence 236, App
29	642.5	31.6	344	3	US-09-990-442-236 Sequence 236, App
30	642.5	31.6	344	3	US-09-991-163-236 Sequence 236, App
31	642.5	31.6	344	3	US-09-993-604-236 Sequence 236, App
32	642.5	31.6	344	3	US-09-990-456-236 Sequence 236, App
33	642.5	31.6	344	3	US-09-989-721-236 Sequence 236, App
34	642.5	31.6	344	3	US-09-992-598-236 Sequence 236, App
35	642.5	31.6	344	3	US-09-989-293A-236 Sequence 236, App
36	642.5	31.6	344	3	US-09-989-735-236 Sequence 236, App
37	642.5	31.6	344	3	US-09-990-444-236 Sequence 236, App
38	642.5	31.6	344	3	US-09-991-181-236 Sequence 236, App
39	642.5	31.6	344	3	US-09-989-730-236 Sequence 236, App
40	642.5	31.6	344	3	US-09-990-436-236 Sequence 236, App
41	642.5	31.6	344	3	US-09-993-687-236 Sequence 236, App
42	642.5	31.6	344	3	US-09-989-734-236 Sequence 236, App
43	642.5	31.6	344	3	US-09-997-653-236 Sequence 236, App
44	642.5	31.6	344	3	US-09-989-724-236 Sequence 236, App
45	642.5	31.6	344	3	US-09-989-728-236 Sequence 236, App

ALIGNMENTS

RESULT 1
US-10-661-430-1
; Sequence 1, Application US/10661430
; Publication No. US20040086955A1
; GENERAL INFORMATION:
; APPLICANT: Cumming, Richard D.
; APPLICANT: Kawai, Ziad S.
; TITLE OF INVENTION: Beta 1, 4-N-ACETYLGLACTOSAMINYLTRANSFERASES,
; TITLE OF INVENTION: NUCLEIC ACIDS AND METHODS OF USE THEREOF
; FILE REFERENCE: 7148.001
; CURRENT APPLICATION NUMBER: US/10/661,430
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 60/411,242
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-661-430-1

Query Match	100.0%	Score 2031;	DB 4;	Length 383;
Best Local Similarity	100.0%	Pred. No. 8.4e-196;		
Matches	383;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1	MAFRLAVARLKSILVLCANVLLVHAMTYKPSLYENTTIGSSTLIADVDAMEAVLGNTA	60	
DB	1	MAFRLAVARLKSILVLCANVLLVHAMTYKPSLYENTTIGSSTLIADVDAMEAVLGNTA	60	
QY	61	STSDLDLTWNSTFSPISVNOQTSFMDIRPLFPDNOQTOCNOQPPHLVGPRIVFIDE	120	
DB	61	STSDLDLTWNSTFSPISVNOQTSFMDIRPLFPDNOQTOCNOQPPHLVGPRIVFIDE	120	
QY	121	PDFKLEKTYPTTAGGCMPCDCVARRVVAIVYRREARLRIMLNHSLAKQOJD	180	
DB	121	PDFKLEKTYPTTAGGCMPCDCVARRVVAIVYRREARLRIMLNHSLAKQOJD	180	
QY	181	YAIFFVEQVANQTFNRGLKMNVGVDVASRLVPMQCFIHHVDVLLPEDDRNLTCPIOPHH	240	
DB	181	YAIFFVEQVANQTFNRGLKMNVGVDVASRLVPMQCFIHHVDVLLPEDDRNLTCPIOPHH	240	
QY	241	MSVAIDKENYKLPYSALFGISALTQDHLKKINGFSNDPFWGSGEDDLATRTSMAGLKV	300	
DB	241	MSVAIDKENYKLPYSALFGISALTQDHLKKINGFSNDPFWGSGEDDLATRTSMAGLKV	300	
QY	301	SRYPQIARVYKMGKSTETATNPVVKCRKYKMGOTRRRTGRLSLKYKLVNLKPLVT	360	
DB	301	SRYPQIARVYKMGKSTETATNPVVKCRKYKMGOTRRRTGRLSLKYKLVNLKPLVT	360	

OY 361 RAVVDLLEKDCRRLRRDPFTCF 383
|||
Db 361 RAVVDLLEKDCRRLRRDPFTCF 383

RESULT 2

US-09-925-301-1072
; Sequence 1072, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1072
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (77)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-301-1072

Query Match 35.4%; Score 718; DB 3; Length 404;
Best Local Similarity 48.9%; Pred. No. 2,9e-63;
Matches 134; Conservative 43; Mismatches 95; Indels 2; Gaps 2;

OY 95 PDNQTLOFCNQTPPHLVGPRIYV-FLDEPDKTLEKIYPTDHAGHGMPKDCVARRVAII 153
|||
Db 80 PAXQGLPYCERSBLLVGPVSVSPSLAEIYERNPRVPEGRRYPACCEPRSRTAII 139
|||
OY 154 VPRYREAHRLIMLHNLSLAKQOLDVYAFIVEQVANOTFNKGKLMNNGYDVASRLYPW 213
|||
Db 140 VPRAREHRLRLYLHHPFLQROQLAYGIYVHOGNGTFNPAKLLNVGREALRDEEM 199
|||
OY 214 QCFIFHDVDLLEPDDNLYTC-PLOPRHMSVAIDKENYKLPYSALFGISALTQDHLKKI 272
|||
Db 200 DCLFLHDVDLLEPDDNLYTC-PLOPRHMSVAIDKENYKLPYSALFGISALTQDHLKKI 259
|||
OY 273 NGFSNDPFGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 332
|||
Db 260 NGFPNEYMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 319
|||
OY 333 QTKRWTRDGLSNLKYKLVNLEKPLYTRAVIDL 366
|||
Db 320 RTQNSWTQDGMNSLTVOQLARELGPLYTNITADI 353
|||

RESULT 3
US-10-132-652-5
; Sequence 5, Application US/10132652
; Publication No. US20030013173A1

; GENERAL INFORMATION:
; APPLICANT: Claussen, Henrik
; APPLICANT: Bennett, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
; FILE REFERENCE: 4305/0E521
; CURRENT APPLICATION NUMBER: US/10/132,652
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US/09/118,464A
; PRIOR FILING DATE: 1998-07-17

; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-652-5

Query Match 35.0%; Score 710; DB 4; Length 393;
Best Local Similarity 48.5%; Pred. No. 1.8e-62;
Matches 133; Conservative 43; Mismatches 96; Indels 2; Gaps 2;

OY 95 PDNQTLOFCNQTPPHLVGPRIYV-FLDEPDKTLEKIYPTDHAGHGMPKDCVARRVAII 153
|||
Db 69 PAXQGLPYCERSBLLVGPVSVSPSLAEIYERNPRVPEGRRYPACCEPRSRTAII 128
|||
OY 154 VPRYREAHRLIMLHNLSLAKQOLDVYAFIVEQVANOTFNKGKLMNNGYDVASRLYPW 213
|||
Db 129 VPRAREHRLRLYLHHPFLQROQLAYGIYVHOGNGTFNPAKLLNVGREALRDEEM 188
|||
OY 214 QCFIFHDVDLLEPDDNLYTC-PLOPRHMSVAIDKENYKLPYSALFGISALTQDHLKKI 272
|||
Db 189 DCLFLHDVDLLEPDDNLYTC-PLOPRHMSVAIDKENYKLPYSALFGISALTQDHLKKI 248
|||
OY 273 NGFSNDPFGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 332
|||
Db 249 NGFPNEYMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 308
|||
OY 333 QTKRWTRDGLSNLKYKLVNLEKPLYTRAVIDL 366
|||
Db 309 RTQNSWTQDGMNSLTVOQLARELGPLYTNITADI 342
|||

RESULT 4

US-11-105-796-5
; Sequence 5, Application US/11105796
; Publication No. US20050181437A1

; GENERAL INFORMATION:
; APPLICANT: CLAUSSEN, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; FILE REFERENCE: 04305/100E521-US1
; CURRENT APPLICATION NUMBER: US/11/105,796
; PRIOR FILING DATE: 2005-04-13
; PRIOR APPLICATION NUMBER: 09/118,464
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-105-796-5

Query Match 35.0%; Score 710; DB 6; Length 393;
Best Local Similarity 48.5%; Pred. No. 1.8e-62;
Matches 133; Conservative 43; Mismatches 96; Indels 2; Gaps 2;

OY 95 PDNQTLOFCNQTPPHLVGPRIYV-FLDEPDKTLEKIYPTDHAGHGMPKDCVARRVAII 153
|||
Db 69 PAXQGLPYCERSBLLVGPVSVSPSLAEIYERNPRVPEGRRYPACCEPRSRTAII 128
|||
OY 154 VPRYREAHRLIMLHNLSLAKQOLDVYAFIVEQVANOTFNKGKLMNNGYDVASRLYPW 213
|||
Db 129 VPRAREHRLRLYLHHPFLQROQLAYGIYVHOGNGTFNPAKLLNVGREALRDEEM 188
|||
OY 214 QCFIFHDVDLLEPDDNLYTC-PLOPRHMSVAIDKENYKLPYSALFGISALTQDHLKKI 272
|||
Db 189 DCLFLHDVDLLEPDDNLYTC-PLOPRHMSVAIDKENYKLPYSALFGISALTQDHLKKI 248
|||
OY 273 NGFSNDPFGMGEGDDDLATRTSMAGLKVSRYPQTQIARYKMIKSTETATNPVNCRYKIMG 332
|||

Db 249 NGFPNEMWGEGEDDDIATRVRLAGMKISRPTSVGHYKMYVGRGDKGNEBNPHRFLLV 308
QY 333 QTKRRWTRDGLSNLKYKLVNLEKPLYTRAVVDL 366
Db 309 RTQNSWTDGMSNLTYQLLARELGPLYNTITADI 342

RESULT 5

US-11-097-143-18555
; Sequence 18555, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 18555
; LENGTH: 403
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-18555

Query Match 34.9%; Score 709; DB 6; Length 403;

Best Local Similarity 42.3%; Pred. No. 2.3e-62;
Matches 150; Conservative 56; Mismatches 129; Indels 20; Gaps 7;

QY 28 IYKISLYENLT---IGSSTLIADVADAMAVGNTASTSDDLLDTWNSTF-SPISEVNQT 83
Db 46 VHKYAHITIGNASSDAGGSE-ASRLPASPLALSKRERQELANGPNSTIRIVIAIANTPT 104
QY 84 SFMEDIRPLT-----FPDNO-----TLQFCNQTPPHLVGPRIVFLDEDFKTLR-KIYP 131
Db 105 SIPQULTRFLGTKKFLPRPKQSTSAALANTCDPDRDGPITPMTTLESIDVIEAEIGP 164
QY 132 DTHAGGCHGPKCVARARVAIIVPRDRARHRIMLAHLSLAAQOQDYAIFIVYQVAN 191
Db 165 LIRPGGAEPENCNAQHVAIVPPRDRVAHLLFLRNHPRLMQORLAIVRFIVEQTNQ 224
QY 192 QTFNRGLKLVNGVYDVA SRLYPWOCFIHVDVLLPEDDRRLYTCPIOPRHMSVAIDKFNK 251
Db 225 KFPNAAAMNIGYLEALKIYQWDCTFIHVDVLLPEDDRRLYTCPIOPRHMSVAIDTLNFR 284
QY 252 LEPYSAIFGGISALTYDHLKKGKFSNDFWGWGEGEDDLATRTSMGLKVSRYPTQIARYK 311
Db 285 LEPYSAIFGGVSAWTEHFOAVNNGFSNDFWGWGEGEDDMSNRLKHANLFIISRYPVNIARYK 344
QY 312 MTKHSTEAHPNPKCRKYKMGQTKRRTDGLSNLKYKLVNLEKPLYTRAVVDL 366
Db 345 MLKHQEKANPK---RYENLQNGMSKIEODGINSIKYSIYSIKOPFTFTWYLAEL 396

RESULT 6

US-10-459-311-7

; Sequence 7, Application US/10459311

; Publication No. US2003020396A1

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; Guegler, Karl J.

; Corley, Neil C.

; Shah, Purvi

; Patterson, Chandra

; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: California

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/459,311

; FILING DATE: 10-Jun-2003

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/373,902

; FILING DATE: 12-Aug-1999

; APPLICATION NUMBER: US/09/055,097

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Cetrone, Michael C.

; REGISTRATION NUMBER: 39,132

; REFERENCE/DOCKET NUMBER: PF-0490 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 855-0555

; TELEFAX: (650) 845-4166

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 362 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 1469908

; SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-459-311-7

Query Match 33.5%; Score 680.5; DB 4; Length 362;

Best Local Similarity 43.9%; Pred. No. 1.5e-59;
Matches 122; Conservative 58; Mismatches 97; Indels 1; Gaps 1;

QY 90 RPILEPDQOTLOFCNQTPPHLVGPRIVFLDEE-DEKTELEKIYPTDTHAGGCHGPKCVARH 148
Db 80 QPRNLIPDSAPSGLCDPSPPLVGPRIVGFSEQPVNLSEVASTNPEVREGGFAPKCKKALQ 139
QY 149 RVAIVPRDRDEARHRIMLAHLSLAAQOQDYAIFIVYQVANQTFNRGLKLVNGVYDVA 208
Db 140 KVAIIIPFRNREHKLKYLWYHPILOQOQDYGYVIVINQDQDEFNRAKLINVFTEAL 199
QY 209 RLYPWOCFIHVDVLLPEDDRRLYTCPIOPRHMSVAIDKFNKLPYSAIFGGISALTYDHL 268
Db 200 KEYVDYDCEVFDVDDIIPMDDRNTYKCYGQPRHLISYMDKFGFRPLYNQYFGGVSAISKEQ 259
QY 269 LKTINGFENDFWGWGEGEDDLATRTSMGLKVSRYPTQIARYKMTKHSATEANPNPKCRY 328
Db 260 FTKINGFNNWGWGEGEDDIYNRLVFKMGKISRDAVIGKCRMTIRSDRRKNEBNPERF 319
QY 329 KIMGQTKRRTDGLSNLKYKLVNLEKPLYTRAVVDL 366

RESULT 10
US-10-466-941-6
; Sequence 6, Application US/10466941
; Publication No. US20040214273A1
; GENERAL INFORMATION:
; APPLICANT: Tatsuji, Seki
; APPLICANT: Kazuhito, Fujiyama
; TITLE OF INVENTION: Method for excretory production of glycoproteins
; TITLE OF INVENTION: having human-type sugar chains
; FILE REFERENCE: J198020401
; CURRENT APPLICATION NUMBER: US/10/466,941
; CURRENT FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: JP200112519
; PRIOR FILING DATE: 2001-01-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-466-941-6

Query Match 32.8%; Score 666; DB 4; Length 385;
Best Local Similarity 35.7%; Pred. No. 4,7e-58;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLCAVLLVHMY-----KIPSLYENLTIGSSTLIADVDAEAVLGNTA----- 60
13 LVAVCALHLGTVLTVYLAGRDSLRLPOL-----VGSTPLQGSNSAAIIGSSGELRG 67
61 -----STSDLLDTWNSTSPISFVNQTSFMEIDRPLIFPDNQTLPQC 103
68 GARPPPLGASSQPRGGSSPVDSGPGPASNLSV-----PVPHTALSLPAC 117
104 NOTPHLVGPTRVFLDEP-DFKTELEKIPDTAGHGMPKDCVARRVAIIIPYDREAH 162
118 PEESEPLVGPMLIEFNMVDELIVAKONPNVGMGRVAPRDCVSPHKVAIIIPFNROEH 177
163 LRIIMLHNLSLAKQOQDYAIFIVEQVANQFNRGKLMNVGVDSRLYPMOCFIHVD 222
178 LKTYLTVLHPVLOQOQDYGIVINQAGDTIFNRKALINVGQELAKDYTCFVPSVD 237
223 LIPEDDRNLVTCPIQPRHMSVAIDKENYKLPYSALFGSISALTQKHLKKNFSDPFMG 282
238 LIPMDHNAVRCFSPRHISVAMDKFSLPVYQVFGVSALSKOQPLTINGFPNNYWG 297
283 GGEEDDLATRTSMAGLKVSRYPTQIARYKMIKSTEATNPVKKCYKIMGQTKRWTRDG 342
298 GGEEDDIFNRVLFVRGMSISRPNAVVGRCMIRHSRDKNEPQRFDRIAHTKETMLSDG 357
343 LSNLKYKLVNLEKPLTYRAVDL 366
358 LNSTLYQVLDVORRPLTYQITVDI 381

RESULT 11
US-10-870-635-6
; Sequence 6, Application US/10870635
; Publication No. US20050143564A1
; GENERAL INFORMATION:
; APPLICANT: Tatsuji, Seki and Kazuhito, Fujiyama
; TITLE OF INVENTION: Method for manufacturing of glycoproteins
; TITLE OF INVENTION: having human-type glycosylation
; FILE REFERENCE: 62,7368
; CURRENT APPLICATION NUMBER: US/10/870,635
; CURRENT FILING DATE: 2004-06-17
; PRIOR APPLICATION NUMBER: US 09/857,651
; PRIOR FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 385

TYPE: PRT
ORGANISM: Homo sapiens
US-10-870-635-6

Query Match 32.8%; Score 666; DB 5; Length 385;
Best Local Similarity 35.7%; Pred. No. 4,7e-58;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLCAVLLVHMY-----KIPSLYENLTIGSSTLIADVDAEAVLGNTA----- 60
13 LVAVCALHLGTVLTVYLAGRDSLRLPOL-----VGSTPLQGSNSAAIIGSSGELRG 67
61 -----STSDLLDTWNSTSPISFVNQTSFMEIDRPLIFPDNQTLPQC 103
68 GARPPPLGASSQPRGGSSPVDSGPGPASNLSV-----PVPHTALSLPAC 117
104 NOTPHLVGPTRVFLDEP-DFKTELEKIPDTAGHGMPKDCVARRVAIIIPYDREAH 162
118 PEESEPLVGPMLIEFNMVDELIVAKONPNVGMGRVAPRDCVSPHKVAIIIPFNROEH 177
163 LRIIMLHNLSLAKQOQDYAIFIVEQVANQFNRGKLMNVGVDSRLYPMOCFIHVD 222
178 LKTYLTVLHPVLOQOQDYGIVINQAGDTIFNRKALINVGQELAKDYTCFVPSVD 237
223 LIPEDDRNLVTCPIQPRHMSVAIDKENYKLPYSALFGSISALTQKHLKKNFSDPFMG 282
238 LIPMDHNAVRCFSPRHISVAMDKFSLPVYQVFGVSALSKOQPLTINGFPNNYWG 297
283 GGEEDDLATRTSMAGLKVSRYPTQIARYKMIKSTEATNPVKKCYKIMGQTKRWTRDG 342
298 GGEEDDIFNRVLFVRGMSISRPNAVVGRCMIRHSRDKNEPQRFDRIAHTKETMLSDG 357
343 LSNLKYKLVNLEKPLTYRAVDL 366
358 LNSTLYQVLDVORRPLTYQITVDI 381

RESULT 12
US-10-761-435A-16
; Sequence 16, Application US/10761435A
; Publication No. US20040241817A1
; GENERAL INFORMATION:
; APPLICANT: Umana, Pablo
; APPLICANT: Bruenker, Peter
; APPLICANT: Ferrara, Claudia
; APPLICANT: Suter, Tobias
; TITLE OF INVENTION: Fusion Constructs and Use of Same to Produce Antibodies with
; TITLE OF INVENTION: Increased Fc Receptor Binding Affinity and Effector Function
; FILE REFERENCE: 1975.0180003
; CURRENT APPLICATION NUMBER: US/10/761,435A
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US 60/441,307
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/491,254
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/495,142
; PRIOR FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: GalT amino acid sequence from pbluegalT
US-10-761-435A-16

Query Match 32.8%; Score 666; DB 5; Length 398;
Best Local Similarity 35.7%; Pred. No. 5e-58;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLCAVLLVHMY-----KIPSLYENLTIGSSTLIADVDAEAVLGNTA----- 60

Db 26 LVAVALHGLVTLVYVYLAGRDLRLPOL-----VGVSTPLGGSSNSAAAIGSSGELRTG 80
QY 61 -----STSDLDLDTMNSTPSPISEVNNQGSFMEIDRPIILFPNOTLQPC 103
Db 81 GARPPPLGASOPRPGGDSFVVDGSGPGASNLTSV-----PVPHTALSRLPAC 130
QY 104 NOTPPIHVGPIRVFLDEP-DFKTLKIKYPTDTHAGGMPKDCVARRHVAIIVRYDREAH 162
Db 131 PESSPLVGVMLIEFMNPVDLELVAKQNPVKKGGRYAPRDCVSPHKVAIIFRRRQEH 190
QY 163 LRTMLNHLSLAKQQLDYAIFIVEQVANOTFNRGKLMNVGVASRLVPMQCFIHHVD 222
Db 191 LKMYLYLHPVLOKQQLDYGVINQAGDTIFNRKLLNVGFQELKDYDTCTCFVSDVD 250
QY 223 LIPEDDRNLITCPIOPRHMSVALDKFENYKLPYSAIFGDISALTTRDHLKIKINGFSNPFMGW 282
Db 251 LIPMNDHNAVRCFSQPRHISVAMDKFGSLPIYQYFGGVALSLSKQOFLTNGFPNNYMGW 310
QY 283 GGEDDLATRTSMAGLKVSRYPTQIARVKMIKISTEATNPNKCRKYMGTTRRTDG 342
Db 311 GGEDDDIFNRLVFRGMSISRPNVVGRCRMIRSRDKNEPFPQFRIRIAHTKETMLSDG 370
QY 343 LSNLTKLVNLKLPYTRAIVVDL 366
Db 371 LNSLTQYVLDVQRYPLTYQTIVDI 394

RESULT 13

US-10-761-435A-20
; Sequence 20, Application US/10761435A
; Publication No. US20040241817A1
; GENERAL INFORMATION:
; APPLICANT: Umana, Pablo
; APPLICANT: Bruenker, Peter
; APPLICANT: Ferrera, Claudia
; APPLICANT: Suter, Tobias
; TITLE OF INVENTION: Fusion Constructs and Use of Same to Produce Antibodies with
; TITLE OF INVENTION: Increased Fc Receptor Binding Affinity and Effector Function
; FILE REFERENCE: 1975.0180003
; CURRENT APPLICATION NUMBER: US/10/761.435A
; CURRENT FILING DATE: 2004-01-22
; PRIOR APPLICATION NUMBER: US 60/441,307
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: US 60/491,254
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: US 60/495,142
; PRIOR FILING DATE: 2003-08-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of ManII-GalT
US-10-761-435A-20

Query Match 32.7%; Score 663.5; DB 5; Length 371;
Best Local Similarity 36.1%; Pred. No. 8e-58; Indels 31; Gaps 6;
Matches 133; Conservative 75; Mismatches 129;

QY 16 VLCAVLLVHAMI-----YKIPSLYENLTIGSSTLIAD-VDAMEAVLGNLTASTDDL 67
Db 14 IFCVIISLVIMLDRGHLDYPRNRRBSPPQGLSMLOEKIDLEHLKLNENNIISIR 73
QY 68 DTNNTSPISSEVNOTSFMEDIRPIILPNDQTLQF-----CNQTPPIHVGPIRVFLD 119
Db 74 D-----SVIMLSSEVED-----GPKSSQSNFSQAGSPACPESPLIVGPMLEFN 119
QY 120 EP-DFKTLEKYPRTTHAGGMPKDCVARRHVAIIVRYDREAHRLIMLHLSLLKQ 178
Db 120 MPVDLELVAKQNPVKKGGRYAPRDCVSPHKVAIIFRRRQEHKLTWLYLHPVLOKQ 179

QY 179 LDYAIFIVEQVANOTFNRGKLMNVGVASRLVPMQCFIHHVDLIPEDDRNLITCPIOP 238
Db 180 LDYGVINQAGDTIFNRKLLNVGFQELKDYDTCTCFVSDVDLIPMNDHNAVRCFSQP 239
QY 239 REMSVAIDENKRLKYSAIFGDISALTTRDHLKIKINGFSNDFMGSGEDDDLATRTSMAGL 298
Db 240 RHISVAMDKFGSLPIYQYFGGVALSLSKQOFLTNGFPNNYMGSGEDDDLATRTSMAGL 299
QY 299 KVSRYPTQIARVKMIKISTEATNPNKCRKYMGTTRRTDGSLNLYKLVNLEKPL 358
Db 300 SISRPNAVVGRCRMIRHSDKKNERPQFRIRIAHTKETMLSDGNSLTQYVLDVQRYEL 359
QY 359 YTRAIVVDL 366
Db 360 YTOITVDI 367

RESULT 14

US-10-713-970-21
; Sequence 21, Application US/10713970
; Publication No. US20050106664A1
; GENERAL INFORMATION:
; APPLICANT: Contreras, Roland
; APPLICANT: Callewaert, Nico L. M.
; APPLICANT: Verwecken, Wouter
; APPLICANT: Kaigorodov, Vladimir
; TITLE OF INVENTION: MODIFICATION OF PROTEIN GLYCOSYLATION IN METHYLOTROPHIC YEAST
; FILE REFERENCE: 17106
; CURRENT APPLICATION NUMBER: US/10/713.970
; CURRENT FILING DATE: 2003-11-14
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-713-970-21

Query Match 32.5%; Score 660.5; DB 5; Length 397;
Best Local Similarity 35.2%; Pred. No. 1.8e-57;
Matches 138; Conservative 73; Mismatches 134; Indels 47; Gaps 5;

QY 7 AVARLSLVLCVLLVHAMIY-----KIPSLYENLTIGSSTLIADVDAMEAVLGN 58
Db 17 SLQACRLVAVCVMHGLVTLVYVYLAGRDLRLPOL-----VGVSTPLGGSSNSAAIGQ 71
QY 59 TA-----STSDLDLDTMNSTPSPISEVNNQGSFMEIDRPIILFP 95
Db 72 SSSELRTGARPPPLGASOPRPGGDSFVVDGSGPGASNLTSV-----PVPHT 121
QY 96 DNQTLQFCNQTPPIHVGPIRVFLDEP-DFKTLKIKYPTDTHAGGMPKDCVARRHVAIIV 154
Db 122 TALSLACRESPSLVGVMLIEFMNPVDLELVAKQNPVKKGGRYAPRDCVSPHKVAIIF 181
QY 155 PYDREAHRLIMLHLSLLAKQQLDYAIFIVEQVANOTFNRGKLMNVGVASRLVPMQ 214
Db 182 PFRNRQEHKLTWLYLHPVLOKQQLDYGVINQAGDTIFNRKLLNVGFQELKDYDT 241
QY 215 CFIEHHVDLIPEDDRNLITCPIOPRHMSVALDKFENYKLPYSAIFGDISALTTRDHLK 274
Db 242 CFVFSVDLIPMNDHNAVRCFSQPRHISVAMDKFGSLPIYQYFGGVALSLSKQOFLT 301
QY 275 FSNDFMGSGEDDDLATRTSMAGLKVSRYPTQIARVKMIKISTEATNPNKCRKYMGT 334
Db 302 FPNNYMGSGEDDDLIFNRLVFRGMSISRPNVVGRCRMIRHSDKKNERPQFRIRIAHT 361
QY 335 KRRWTRDGLSNLYKLVNLEKLPYTRAIVVDL 366
Db 362 KETMLSDGNSLTQYVLDVQRYPLTYQTIVDI 393

RESULT 15

US-10-287-226-82

```

Sequence 82, Application US/10287226
Publication No. US20040086875A1
GENERAL INFORMATION:
APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Berghs, Constance,
APPLICANT: Boldog, Ferenc,
APPLICANT: Burgeess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eissen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gangolli, Esna A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekuda, Ramesh,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Malyankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigaru, Muralidhara,
APPLICANT: Paturajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vermet, Corine A.M.,
APPLICANT: Zernusen, Bryan D.,
APPLICANT: Zhong, Wei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/334,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: Curoseglist version 0.1
SEQ ID NO 82
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
US-10-287-226-82

```

Query Match	Similarity	32.5%	Score 660;	DB 4;	Length 372;
Best Local	Similarity	37.5%;	Pred. No. 1.8e-57;		
Matches 137;	Conservative	70;	Mismatches 132;	Indels 26;	Gaps 7
QY	12	KSLLVLCVALLLVNAMIYKIPSLYEYNTLTGSS-TLLADVDAAMEAVLNGTASTSDDLLDTW	70		
Db	14	KAVLLLCILHRLVAVILY-FDVYQHLIAFFSRFSARGBANHLHPAASSSSSSSSCRPNVA	72		
QY	71	NSTSSPISEVNVQTSFMEDIRILFPRPDNLOLCNQTNPRLVGPRLVFLDEPRFKLTKIY	130		
Db	73	TASSSGRLPEVSA-----LPSPATPLPPCDSPPGVIGRLILFTSP--MPLERVN	122		
QY	131	---DPTAGGSHGMPKDCVARRHVAIIIVPYRDREAHRLIMLHNLSLAKOOLDUYAFIVE	187		
Db	123	RENGVLMGGGYTPRDCPRQTVAVIIPFRRENHLRWLYHNLHILAKQRLRYGVYVIN	182		
QY	188	QVANQTNRRGKLVNNGY-----DVASRLYWCQSFIFHDVDLLREDDENLYTCPIQRPM	241		
Db	183	QHGEDTFRRAKLVNNGVLEALKEBAA-----YDFEIFSDVVLVPMDDENLVRCGQPRHF	237		
QY	242	SVALDKFNFKLPSYSAIFGGSIALTKRDHLKKINGNSPDMGWMGGEDDDLATYSMAGLKVS	301		
Db	238	AIADMDKGFRLPYLYGYFGVSGSLKAQFLRINGSPNENWGMGGEDDDLFNRIISLGKMTIS	297		
QY	302	RYPTQIARYKMIKSTEAATPVNKCRRYKIMQTKRRWTRDGI.SNLKLYLVLELKPPLYTR	361		
Db	298	RPIRIRIGRYMKIKDRDKHNEPNPQRFYKIQNTKLYTMKRGDGISVRYQVLVEVSRQPLFTN	357		
QY	362	AVVDL 366			
Db	358	ITVDI 362			

Search completed: January 6, 2006, 15:33:53
Job time : 116 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:15:29 ; Search time 39 Seconds
(without alignments)
944.898 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRIHAVRLKSLIVLCAN.....VDLEKDCRRLRDPPTCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	667.5	32.9	399	2	A33396 beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - mouse
2	666	32.8	397	2	A01030 beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - mouse
3	648.5	31.9	402	2	S05018 beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - mouse
4	569.5	28.0	224	2	T46511 hypothetical protein
5	523.5	25.8	490	2	A55141 G1CNAc beta-1,4-N-hypothetical prote
6	487.5	24.0	370	2	T26091 galactosyltransferase
7	259	12.8	289	2	S40716 beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - mouse
8	244.5	12.0	194	2	PC1175 chitinase
9	114	5.6	110	2	A39567 chitinase
10	107.5	5.3	105	2	S76557 chitinase
11	105.5	5.2	715	2	T25233 chitinase
12	103	5.1	1103	2	H82884 chitinase
13	99.5	4.9	623	2	T42345 chitinase
14	99.5	4.9	624	2	T42347 chitinase
15	99.5	4.9	626	2	T42346 chitinase
16	98	4.7	1175	2	S39551 chitinase
17	96	4.7	579	2	T31549 chitinase
18	94	4.6	368	2	S49712 chitinase
19	93.5	4.6	585	2	B69318 chitinase
20	93.5	4.6	1187	2	F86422 chitinase
21	93	4.6	1510	2	A64120 chitinase
22	92.5	4.6	363	2	A83470 chitinase
23	90.5	4.5	620	2	F83169 chitinase
24	90.5	4.5	876	1	A57988 chitinase
25	90	4.4	431	2	C88957 chitinase
26	89.5	4.4	443	2	AF1901 chitinase
27	89	4.4	563	2	A88515 chitinase
28	89	4.4	612	2	T42243 chitinase
29	88.5	4.4	647	2	C87440 chitinase

30	88.5	4.4	1164	2	F72166 A25R protein - var
31	88	4.3	358	2	T29492 hypothetical prote
32	88	4.3	1785	2	S53976 probable membrane
33	87.5	4.3	276	2	AG2532 hypothetical prote
34	87.5	4.3	314	2	G90365 conserved hypotet
35	87.5	4.3	509	2	AE1295 Bitfunctional phosp
36	87.5	4.3	776	2	T15411 hypothetical prote
37	87.5	4.3	872	2	E71852 valine-tRNA ligase
38	87.5	4.3	1156	2	T37411 RNA polymerase sub
39	87.5	4.3	1164	1	RMV28T DNA-directed RNA p
40	87.5	4.3	1164	1	T28566 DNA-directed RNA p
41	87.5	4.3	2083	2	T42721 CRP-ductin-alpha p
42	87	4.3	336	2	A97168 glycosyltransferase
43	87	4.3	661	2	S37590 beta-fructofuranos
44	87	4.3	736	2	T40904 pome specific hyp
45	87	4.3	1215	2	S60904 hypothetical prote

ALIGNMENTS

RESULT 1
A33396
beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - mouse
N:Alternate names: glycoprotein 4-beta-galactosyltransferase; UDP-beta-1,4-galactosyltrar
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #ext_change 09-Jul-2004
C:Accession: A33396, A28799, JX0041, I49531, I49532
R:Hollis, G.F.; Douglas, J.G.; Shaper, N.L.; Shaper, J.H.; Stafford-Hollis, J.M.; Evans, R.; Biophys. Res. Commun. 162, 1069-1075, 1989
A:Title: Genomic structure of murine beta-1,4-galactosyltransferase.
A:Reference number: A33396, MUID:89350913, PMID:2504153
A:Accession: A33396
A:Molecule type: DNA
A:Residues: 1-399 <HOL>
A:Cross-references: UNIPROT:P15535; UNIPARC:UPI0000022B9E; GB:M27922; NID:G341810; PIDN:f
A:Note: the authors did not translate the codon GTT for residue 139
R:Shaper, N.L.; Hollis, G.F.; Douglas, J.G.; Kirsch, I.R.; Shaper, J.H.
J. Biol. Chem. 263, 10420-10428, 1988
A:Title: Characterization of the full length cDNA for murine beta-1,4-galactosyltransfer
A:Reference number: A28799, MUID:88273147, PMID:3134348
A:Accession: A28799
A:Molecule type: mRNA
A:Residues: 1-399 <SHA>
A:Cross-references: UNIPARC:UPI0000022B9E; EMBL:J03880; NID:G192185; PIDN:AAA37297.1; PII
A:Note: it is uncertain whether Met-1 or Met-14 is the initiator
R:Nakazawa, K.; Ando, T.; Kimura, T.; Narimatsu, H.
J. Biochem. 104, 165-168, 1988
A:Title: Cloning and sequencing of a full-length cDNA of mouse N-acetylglucosamine (beta)
A:Reference number: JX0041, MUID:89033997, PMID:3141392
A:Accession: JX0041
A:Molecule type: mRNA
A:Residues: 1-359 <NAK>
A:Cross-references: UNIPARC:UPI0000022B9E; GB:D00314; NID:G220340; PIDN:BA00216.1; PID:f
R:Shaper, J.H.; Hollis, G.F.; Shaper, N.L.
Biochimie 70, 1683-1688, 1988
A:Title: Evidence for two forms of murine beta-1,4-galactosyltransferase based on cloning
A:Reference number: I49531, MUID:89207607, PMID:3149531
A:Accession: I49531
A:Molecule type: mRNA
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Accession: I49531
A:Accession: I49532
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 14-63 <RE2>
A:Cross-references: UNIPARC:UPI000000492; GB:M36289; NID:G192188; PIDN:AAA37295.1; PID:f
C:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase; transmembrane protein
F:25-44/Domain: transmembrane #status predicted <TM>
F:113/Binding site: carbohydrate (Aen) (covalent) #status predicted

	Query March	32.9%;	Score 667.5;	DB 2;	Length 399;
	Best Local Similarity	45.7%;	Pred. 3.1e-50;		
	Matches 123;	Conservative 52;	Mismatches 93;	Indels 1;	Gaps 1;
Qy	99	TLQFCNQTPHLVGPPIRV-FLIDEPDFKLEKITYDTHAGHGMPKDCYAKRRRAVIAIYPR	157		
Db	127	SLPACPESSPLLVPBMIDFNIAVDLELLAKNPEIKTGGYSPKDCVSPHKAVIIIPFR	186		
Qy	158	DREAHLEIMLHNLISLAKQQLDYAIFIVEQVAAQTENRGLNMVGYDVASRLPWCOCI	217		
Db	187	NRGHILAKWMLYYLPHIILQROQDDYGIYINQAGPTMFNRAKLNIIGFQALKQDYNCFV	246		
Qy	218	FHDVVDLPEDDRNLITCPIDPRHMSVAIDKNYKLLPFSALIGGIALTKLTKKINFSN	277		
Db	247	FSDVDLLPMDRRNAYRCFSQPRHISVAMDKGFSLPYVQFGVGSALSKOOFIAINFSN	306		
Qy	278	DFPMGGEDEDDLATRTSMAGLKVSRYPQIARLYMKIKHSTEATIPVANKCRKYLWGQTKRR	337		
Db	307	NYWMGGEDEDDIIFRLVYHKMSISRPAVAVGRCMIRHSRDKKNPEPQDRDRIAHTKET	366		
Qy	338	WTRDGLSLTKYLVNLEKPLIYRAVVDL	366		
Db	367	MRFDGLNSLTYKVLVDQRYPLYQTQIVDI	395		

RESULT 2

J01030

beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) - human

N/Alternate names: N-acetylglucosaminide (beta 1-4)-galactosyltransferase, UDP-beta-1,4-galactose 4-epimerase, lactose synthase (EC 2.4.1.22) protein A; N-acetylglucosamine synthase (EC 2.4.1.22)

C/Species: Homo sapiens (man)

C/Date: 30-Sep-1991 #sequence revision 30-Sep-1991 #text change 09-Jul-2004

C/Accession: J01030; A31353; S13792; S37717; A24251; A24229; S12047

R/Mengle-Gaw, L.; McCoy-Haman, M.F.; Tiemeier, D.C.

Biochem. Biophys. Res. Commun. 176, 1269-1276, 1991

A>Title: Genomic structure and expression of human beta-1,4-galactosyltransferase.

A/Reference number: J01030; MUID:91248214; PMID:1903938

A/Accession: J01030

A/Molecule type: DNA

A/Residues: 1-397 <MEN>

A/Cross-references: UNIPROT:P15291; UNIPARC:PI000012671C; GB:M70427; GB:M70431

R/Maeri, K.A.; Appert, H.E.; Fukuda, M.N

Biochem. Biophys. Res. Commun. 157, 657-663, 1988

A>Title: Identification of the full-length coding sequence for human galactosyltransferase

A/Reference number: A31353; MUID:89076299; PMID:3144273

A/Accession: A31353

A/Molecule type: mRNA

A/Residues: 1-211, 'GIT', 212-397 <MAS>

A/Cross-references: UNIPARC:PI0000158706; GB:M2921

R/Matzele, G.; Berger, E.G.

Nucleic Acids Res. 18, 7174, 1990

A>Title: Near identity of HeLa cell galactosyltransferase with the human placental enzyme

A/Reference number: S13792; MUID:91088335; PMID:2124683

A/Accession: S13792

A>Status: Preliminary

A/Molecule type: mRNA

A/Residues: 1-10, 'S', 11-397 <MAT>

A/Cross-references: UNIPARC:PI000002022E; EMBL:X55415; NID:G32057; PIDN:CAA39073.1; PITN:U01jima, T.; Uemura, M.; Nozawa, S.; Narimatsu, H.

ancer Res. 52, 6158-6163, 1992

A>Title: Complementary DNA cloning for galactosyltransferase associated with tumor and c

A/Reference number: S37717; MUID:93046112; PMID:1384956

A/Accession: S37717

A>Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: mRNA

A/Residues: 1-9, 'RS', 11-33, 'R', 35-397 <UED>

A/Cross-references: UNIPARC:PI0000158549; EMBL:X13223; NID:G34989; PIDN:CAA31611.1; PITN:R/Albert, H.E.; Rutherford, T.J.; Tarr, G.E.; Wiest, J.S.; Thomford, N.R.; McCorquodale, Biochem. Biophys. Res. Commun. 139, 163-168, 1986

A>Title: Isolation of a cDNA coding for human galactosyltransferase.

A/Reference number: A24251; MUID:87025694; PMID:3094506

A:Accession: A24251
A:Molecule type: mRNA
A:Residues: 137-397 <APP>
A:Cross-references: UNIPARC:UPI000016A9F4, GB:M13701, NID:g183659, PID:AAA35935.1, PID:g183659
R:Appert, H.E.; Rutheford, T.J.; Tair, G.E.; Thomford, N.R.; McCordudale, D.J.
Biochem. Biophys. Res. Commun. 138, 224-229, 1986
A:Reference number: A24229; MUID:86295672; PMID:3091013
A:Accession: A24229
A:Molecule type: Protein
A:Residues: 78-93 <AP2>
A:Cross-references: UNIPARC:UPI0000179825
A:Experimental source: milk
R:Aoki, D.; Appert, H.E.; Johnson, D.; Wong, S.S.; Fukuda, M.N.
EMBO J. 9, 3177-3178, 1990
A:Title: Analysis of the substrate binding sites of human galactosyltransferase by protei
A:Reference number: S12047; MUID:91006010; PMID:2120039
A:Accession: S12047
A:Status: preliminary
A:Molecule type: protein
A:Residues: 273-325 <AOK>
A:Cross-references: UNIPARC:UPI0000179826
C:Comment: This enzyme catalyzes the transfer of galactose from an activated UDP-galactose
C:Genetics:
A:Gene: GDB:GCTB2
A:Cross-references: GDB:119981; OMIM:137060
A:Map position: 9p21-9p13
A:Introns: 136/3, 215/3, 278/3, 319/3, 355/3
C:Keywords: glycosyltransferase, hexosyltransferase, lactose biosynthesis, transmembrane

```

Query Match Summary          32.8%; Score 666; DB 2; Length 397;
Best Local Similarity       35.7%; Pred. No. 4,1e-50;
Matches 137; Conservative   73; Mismatches 128; Indels 46; Gaps 5;

QY      14 LVLCAVLLVHAMTY-----KIPSLYENLTIGSSTLIADVDAEAVLGNTA-----60
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      25 LVAVCALHLGYTLVYYLAGRDSLRLPOL-----VGSTPLQGSSNSAAALGQSSGELRTG 79
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY      61 -----STSDDLPTWNSTFSPISSEVNQTSFMEDIRPILFPNQTLQFC 103
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      80 GARPPPLPGASSPRRGDSSPVDSGGPSANLSLTV-----PVPHTTALSIPAC 129
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY     104 NOTPHLVGPPIRFVFLDEP-DFTKLEKIYVDTHAGHGMPDCVARRHVALIIVPRDEAH 162
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB     130 PEESFLVGPMLIERNMPVDLELVAQNPNVXMGGVAPRDCCVSPHKVALIIIFRNQEH 189
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY     163 LRIMLNHNLSIAAKOOLDPAIFIVEQVANOTENFRGLMYDVASRLYPMQCFIFFHDVD 222
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB     190 LKMYLYLHPVLRQROLDYGIVINQAGDTIFFRATLNVNGPEALKKDVIYTCVFSDVD 249
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY     223 LIPEDDRNLVCTPIOPRHMSVAIDKENNYLKLPSAITGGISALTDKHLKKIKINGSNDPWGW 282
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB     250 ILPMNDHNARVCFSPRIHTSVAMDKFGSLPYVQYGVGSALSKOOFLTINGEPNNYWGM 309
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY     283 GGEDDDLARTSMAGLKVSRYPTQIARYKNMIRKISTEATNEVNRKCRXYKIMGQTKRRMRPDG 342
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB     310 GGEDDDIIFRLVYFRGMSISRPNNAVGRCRMIRSRKQKEPNQRPFRIATHTETMTSDG 369
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY     343 ISNLKYKLVNLEKPLYTRAVVDL 366
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB     370 INSLTYQVLDVORYPLTYQTIVDI 393
           : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 3
S05018
beta-N-acetylglucosaminylalyclopeptide beta-1,4-galactosyltransferase (EC 2.4.1.38) precursor
N/Alternate names: beta 1,4-galactosyltransferase
N/Contains: beta-N-acetylglucosaminylalyclopeptide beta-1,4-galactosyltransferase (EC 2.4.
C/Species: Bos primigenius taurus (cattle)
C/Date: 07-Sep-1990 #sequence revision 07-Sep-1990 #text change 09-Jul-2004
C/Accession: S05018; A55077; A24148; A25129; A39076; A37809; B35077
R/d/Agoscaro, C.; Bendak, B.; Tropak, M.
Eur. J. Biochem. 183, 211-217, 1989
File: Cloning of cDNA encoding the membrane-bound form of bovine beta 1,4-galactosyltr

```


A:Reference number: S05018; MUID:89325338; PMID:2502398
A:Molecule type: mRNA
A:Residues: 1-402 <DDG>
A:Cross-references: UNIPROT:P08037, UNIPARC:UPI0000168534; EMBL:X14558; NID:g381; PIDN:C
A:Experimental source: liver
A>Note: part of this sequence was confirmed by protein sequencing
R:Russo, R.N.; Shaper, N.L.; Shaper, J.H.
J. Biol. Chem. 265, 3324-3331, 1990
A>Title: Bovine beta1->4-galactosyltransferase: two sets of mRNA transcripts encode two
each the short and the long forms of the enzyme are type II membrane-bound glycoproteins
A:Reference number: A35077; MUID:90153986; PMID:2105947
A:Accession: A35077
A:Molecule type: mRNA
A:Residues: 1-157,'V','I',159-165 <RUS>
A:Cross-references: UNIPARC:UPI000016C319; GB:J05217; NID:g163125; PIDN:AAA30559.1; PID:
R:Naimatsu, H.; Sinha, S.; Brew, K.; Okayama, H.; Qasba, P.K.
Proc. Natl. Acad. Sci. U.S.A. 83, 4720-4724, 1986
A>Title: Cloning and sequencing of cDNA of bovine N-acetylglucosamine (beta-1-4)galactose
A:Reference number: A24148; MUID:86259693; PMID:3014508
A:Accession: A24148
A:Molecule type: mRNA
A:Residues: 74-157,'V','I',159-163,'E',165-186,'P',188-204,'IL',207-255,257-264,'Q',266-402
A:Cross-references: UNIPARC:UPI0000179827; GB:M13569; NID:g163407; PIDN:AAA30659.1; PID:
A:Experimental source: mammary gland
R:Shaper, N.L.; Shaper, J.H.; Weuth, J.L.; Fox, J.L.; Chang, H.; Kirsch, I.R.; Hollis, G
Proc. Natl. Acad. Sci. U.S.A. 83, 1573-1577, 1986
A>Title: Bovine galactosyltransferase: identification of a clone by direct immunological
A:Reference number: A25129; MUID:86149345; PMID:2419911
A:Accession: A25129
A:Molecule type: mRNA
A:Residues: 69-157,'V','I',159-186,'P',188-204,'IL',207-281,'U',283-402 <SHA>
A:Cross-references: UNIPARC:UPI000016C30F; GB:M13214; NID:g163073; PIDN:AAA30534.1; PID:
A:Experimental source: MDBK cells
R:Yadav, S.P.; Brew, K.
J. Biol. Chem. 266, 698-703, 1991
A>Title: Structure and function in galactosyltransferase. Sequence locations of alpha-1a
A:Reference number: A39076; MUID:91093259; PMID:1898734
A:Accession: A39076
A:Molecule type: protein
A:Residues: 81-95,171-179,'X',192-197,217-222,225-243,266-269,337-342,363,'X',365-373
A:Cross-references: UNIPARC:UPI0000179828; UNIPARC:UPI0000179829; UNIPARC:UPI000017982A,
82F; UNIPARC:UPI0000179830
R:Yadav, S.; Brew, K.
J. Biol. Chem. 265, 14163-14169, 1990
A>Title: Identification of a region of UDP-galactose-N-acetylglucosamine beta4-galactosy
A:Reference number: A37809; MUID:90354395; PMID:2117606
A:Accession: A37809
A:Molecule type: protein
A:Residues: 81-89,'X',91-95,171-180,190-198,'YL',201,209-216,229-238,'H',240-244,351-358
A:Cross-references: UNIPARC:UPI0000179831; UNIPARC:UPI0000179832; UNIPARC:UPI0000179833,
C:Genetics
A:Gene: gaitr
C:Keywords: alternative initiators; glycoprotein; glycosyltransferase; hexosyltransferase
F:1-18/Domain: signal sequence #status predicted <SIG>
F:14-402/Product: beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase,
F:19-402/Product: beta-N-acetylglucosaminylglycopeptide beta-1,4-galactosyltransferase,
F:25-44/Domain: transmembrane #status predicted <TM>
F:79-250/Region: alpha-lactalblumin binding #status predicted
F:341-351/Region: UDP-galactose binding #status predicted
F:90,117/Binding site: carbohydrate (asn) (covalent) #status experimental
F:134-247/Bisulfide bonds: #status experimental

Query Match 31.9%; Score 648.5; DB 2; Length 402;
Best Local Similarity 43.0%; Pred. No. 1,4e+48;
Matches 119; Conservative 55; Mismatches 102; Indels 1; Gaps 1;

Dd 91 FILEPFNOTIOFCNQTPRHVLGPVRFLDEB-DFTKEIKTYPDTHAGHGMPKDVCVARHR 149
:::|||||:::|||||:::|||||:::|||||:::|||||
Db 122 FVPSTTRSLACPEESPLLVGMPIEBPNIEVDKLIEQNPKVKLGGRYVPMDCISPHK 181
|||:::|||:::|||||:::|||||:::|||||:::|||||:::|||||
Oy 150 VAIIVPRDRARLRIMAHNIHSLLAKOULDIAFIYEQVANOTFRNGKLTANVGVDVASR 209
|||:::|||:::|||||:::|||||:::|||||:::|||||:::|||||

Db 182 VAIITLFFNRQGEHLKXWLYLHPWQROQLDGIYVINGQSGEMFNRAKLINVFKEALK 241

Qy 210 LYPWOCPIFHDDVLLPEDDRNLVTCPTQPRHMSVAIDKENYKLPYSALFGISALTQDHL 269

Db 242 DYDYCCFFVSVDVLLPMDNDHTYRCFSQPRHISIVAMDKGFSPLPYQYFGVSALSKQF 301

Qy 270 KKNESFSDPFWMGGEEDDLATRTSMAGLKVSRYPQIARVKNILKSTTEATNPVNCGRYK 329

Db 302 LSNFPPNNYWMGGEEDDIYNRLAFRGMSVSPNPAVIGCRMIRHSRDKKNFNPORFD 361

Qy 330 IMGOTRRRTRDGLSNLKYKLVNLELKPPLYTRAVVDL 366

Db 362 RIATHKETMLSDGLNSLTYTMVLEVOGRYPLTYKTIVDI 398

RESULT 4

T46511

hypothetical protein DKFZp586M2424.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #ext_change 09-Jul-2004

C:Accession: T46511

R:Duesterhoef, A.; Lauber, J.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223036

A:Accession: T46511

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-224 <AAA>

A:Cross-references: UNIPROT:060909; UNIPARC:UPI000016ACCC; EMBL:ALJ3647

A:Experimental source: adult uterus; clone DKFZp586M2424

A:Genetics:

A>Note: DKFZp586M2424.1

Query Match 28.0%; Score 569.5; DB 2; Length 224;

Best Local Similarity 46.1%; Pred. No. 4,8e-42;

Matches 101; Conservative 44; Mismatches 63; Indels 11; Gaps 2;

Qy 154 VPYRDREAHRLIMLNHLSLAKQQLDVAIFIVEQVANOTFNKGLMNVGY-----DVA 207

Db 1 IPRFRHREHLRLWHLHYLHPIILRRQRLRYGVYVINGHGEDTFENRAKLINVGFLALKEADA 60

Qy 208 SRLVWOCPIFHDDVLLPEDDRNLVTCPTQPRHMSVAIDKENYKLPYSALFGISALTQD 267

Db 61 -----YDFEITFSDVDLVPMDDENLYRCGDQPRFHALAMDGFGRLLPYAGYFGVSGLSKA 115

Qy 268 HLKKNINGSNDPMWGGEDDDLATRTSMAGLKVSRYPQIARVKNILKSTTEATNPVNCR 327

Db 116 QPLRINFPNENWMGGEEDDIFNRLSLTGKMSKRDIDIRIGRFRMITHKDRDKHNEPAPQR 175

Qy 328 YKINGQTKRRWTRDGLSNLKYKLVNLELKPPLYTRAVVDL 366

Db 176 FTKIGNTKLTMKRDGIGSVRYQVLEVSRLPFTNITVDI 214

RESULT 5

A55141

GLNAC beta-1,4-N-acetylglucosaminyltransferase (BC 2.4.1.-) - great pond snail

C:Species: Lymnaea stagnalis (great pond snail)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #ext_change 09-Jul-2004

C:Accession: A55141

R:Bakker, H.; Agteleberg, M.; Van Tetering, A.; Koelmaan, C.A.M.; Van den Bijnden, D.H.; \

J. Biol. Chem. 269, 30326-30333, 1994

A:Title: A Lymnaea stagnalis gene, with sequence similarity to that of mammalian beta1 -

A:Reference number: A55141; MUID:95074032; PMID:7527028

A:Accession: A55141

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-490 <BAK>

A:Cross-references: UNIPROT:009323; UNIPARC:UPI000012676B; GB:X80228; NID:9620087; PID:9620087

C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 25.8%; Score 523.5; DB 2; Length 490;


```
QY 185 IVEQVANQTFNGKLMNVGYDVASRLYPWQCFIFHDVLLPEDDRNLVTCPIQPR----- 239
  ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 1 VTNQAGDEFEFGAKLLNVGFTALMEYDTCTCFVSDVLI--QXXQTLCTQCGRPDPNG 58

QY 240 ---HMSVALDKENYKLPYSAIFGGIS-----ALTKDHLLKINGFSNDFW 280
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 59 XQEHLOVLOPT---KAPF-CLHGXIQIGFYPTISIMSVAEORTIHEDH-----GFPNNYW 109

QY 281 GNGGDDDLATRTSMAGLKVSRYPTQIARYKMIKSTEXTNVNCKRYKIMQTKRRMTR 340
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 110 GNGGDDDLTYNRLVFKMGWISRPDAVIGKCRMI RSHRTKNEPNDRFRIATRETWS 169

QY 341 DGLSNLKY 348
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 170 DGLKSLTY 177

RESULT 9
A39567
thyroxine-binding globulin precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Feb-1992 #sequence_revision 13-Mar-1997 #text_change 05-Oct-2004
C:Accession: A39567; 153286
R:Imamura, S.; Mori, Y.; Murata, Y.; Yamamori, I.; Miura, Y.; Oiso, Y.; Seo, H.; Matsui,
Biochemistry 30, 5406-5411, 1991
A>Title: Molecular cloning and primary structure of rat thyroxine-binding globulin.
A:Reference number: A39567; PMID:91242454; PMID:1903654
A:Accession: A39567
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 10-418 <IM2>
A:Cross-References: UNIPROT:P35577; UNIPARC:UPI0000136DZF; GB:M63991; NID:9707159; PIDN:
R:Tani, Y.; Mori, Y.; Miura, Y.; Okamoto, H.; Inagaki, A.; Saito, H.; Oiso, Y.
Biochemistry 135, 2731-2736, 1994
A>Title: Molecular cloning of the rat thyroxine-binding globulin gene and analysis of its
A:Reference number: 153286; PMID:95080147; PMID:7988464
A:Accession: 153286
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-23 <RES>
A:Cross-References: UNIPARC:UPI0000082B23; GB:S75724; NID:913170; PIDN:AAB32676.1; PID:
A:Experimental source: strain Wistar
C:Superfamily: serpin

Query Match 5.6%; Score 114; DB 2; Length 418;
Best Local Similarity 22.7%; Pred. No. 0.059;
Matches 94; Conservative 62; Mismatches 150; Indels 108; Gaps 20;

QY 25 HAMIKIPIS-----LYENLTIGSSTL-----IADVAMEAVL--GNTASTSDLDLT 69
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 39 NATILKMEPSINDFAFLRLKLSVENPDLNIFPSPISIALAMLSFGSGSSTQTOILEV 98

QY 70 --WNSTFPSISEVNOTSFEMEDIRPILFPDNO--TLQFCNQTPPHLVG---PIRVLEDEP 122
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 99 LGFNLIDTDPVKELQ--GFGHLICSLNFPNNELELQGNNAV---FIGQQLKPLAKFLD--D 152

QY 123 FTTLEKI-----YEDTHAGHGMPKDCVARRHVAIIYVRREARLRMLNHLHLAK 176
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 153 VETLVEVEVSTDFSVSAQHEINSYKQTKGKLVGLIDLIKINIIMLVNYIHFKAQ 212

QY 177 QQLDAVAFIVEQVANQTFNGKLMNVGYDVASRLYPWQCFIFHDV-----LLPEDDR 229
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 213 WANPRVSTKESSN--FVVDKSTTYQVPMHGLEOY---YHYVDVLANCTVLQMDYSA 266

QY 230 NLYTCPIQPR--HM-----SVALDKENY-----KLPYSAIFGISAALTYD 267
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 267 NALALFVLEKEGMEVWEAAMSSKTLKKNVHLQKGWELFVPKFSISATYDLSGTQKM 326

QY 268 HLKIKNGFNDPMWGNGEDDLATRTSMAGLKVSRYPTQIARYKMIKSTEXTNVNCKR 327
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 327 GMRDPAFASADFPG-----ITKDNGLKLS-----YAFHKAVALHIG----- 362
```

```
QY 328 YKIMGQTKRWRTRDGLSNLKYKLVNLELKPITYRAVVD-----LLEKDCREL 375
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 363 -----EGTKEASPEAGSLDQPEVAPLHAVALRLDRFILMLLEKRTSVL 407

RESULT 10
S76557
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) large chain [similarity]
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Sep-2000
C:Accession: S76557
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yasuda,
DNA Res. 3, 109-136, 1996
A>Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
S.
A:Reference number: S74322; PMID:97061201; PMID:8905231
A:Accession: S76557
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1105 <KAN>
A:Cross-References: UNIPARC:UPI0000164C78; EMBL:D64002; GB:AB001339; NID:91001612; PIDN:
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Start codon: GTG
C:Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain; biotin C
C:Keywords: ligase
F:33-1086/Domain: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain homolo
F:33-494/Domain: biotin carboxylase homology <BC1>
F:584-1024/Domain: biotin carboxylase homology <BC2>

Query Match 5.3%; Score 107.5; DB 2; Length 1105;
Best Local Similarity 22.6%; Pred. No. 0.85;
Matches 84; Conservative 45; Mismatches 115; Indels 127; Gaps 21;

QY 25 HAMIKIPISLYENLTIGSS--TLADVAMEAVLNGTASTSDLDLTWNSTFSPISSEVQ 82
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 607 HAAFLSDAGYETITVNSNPETVSTDY-----TSDRLX-----FELTK--- 646

QY 83 TSFEMEDIRPILFPDNOT--LQFCNQTPPHLVGPIRVLEDEDFTELEKIY---DPTNAG 136
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 647 ----EDVNIITEAEKPVGIIIFQGGQTPKLAVPQKYLNSPDCVQKIKWGTSPDS--- 699

QY 137 GHGMPKDCVARRHVAIIYVRDREARLRMLNHLHLAKQ-----QDYAFIYEQVAN 191
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 700 -----IDYADREFEKI-----LHELEISQPNGIARDYESSRV--VAN 737

QY 192 QTFNGKLMNVGYDVASRLYPWQCFIFHDVLLPEDDR-----NLYTCPIQPRHMSVAIDK 247
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 738 R-----ISYPPVVR--PSYVLGGRAMEIVSDELEKRYMYAVQILEPH--PILIDK 785

QY 248 FNYKLPYSAIFGISAALTQDLKIKINGFSNDFWGNGEDDLATRTSMAGLKVSRYPTQI 307
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 786 F-----LENAIEVDVSLDSTGKVIIG-----SIHEIEBAGI----- 819

QY 308 ARYKMIKSTEA--TNPNNCKRYKIMQTKRRWRD-----GLSNLKY-----KLVN 352
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 820 -----HSGSAGCIPYTSLSDNVL--TTIROWTEQLARALNVGLMNIQYAVGDQYVI 871

QY 353 LELKPLYTRAV 363
  |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 872 LEANPRASRTV 882

RESULT 11
T52333
hypothetical protein T24D1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T52333
R:Cummings, P.
submitted to the EMBL Data Library, October 1996
```

A:Reference number: Z20001
A:Accession: T25233
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-715 <MIL>
A:Cross-references: UNIPROT:002330; UNIPARC:UPI000008000C; EMBL:Z81131; PIDD:CA803422.1;
A:Experimental source: clone T24D1
C:Genetic:
A:Gene: CESP:T24D1.1
A:Map position: 1
A:Introns: 3/3; 99/3; 210/2; 371/3; 514/2; 629/3; 655/3

Query Match 5.2%; Score 105.5; DB 2; Length 715;
Best Local Similarity 25.0%; Pred. No. 0.69;

Matches 53; Conservative 21; Mismatches 69; Indels 69; Gaps 8;

QY 152 IIVPYDREAHRLIMLHLSLAKQOLDYAI-----FI 185
DB 467 MIMPLRGRAIFARFAQHLKSCICARGGDDLAVALITVYSSSEDEMENRETIEMLRASFLP 526
QY 186 --VEQVANQTFNRGKLMNVGYDVASRLYPWQCFIPHDVDDLPRDRLNLTGPIQPRHMSV 243
DB 527 VTIEMGDVFSFGVALMRG--AETLPANALLFTDVM-----LFTCDALKIKSN 576
QY 244 AI-----DKFNY--KLPSAIFG-GISALTRDHLKKIN 273
DB 577 TLINAGIPIPIVSEFSHESMSNDKLLADAFHYGRGRGYFRHFGYGLAMVYKADLMDVG 636
QY 274 GFSNDFWVGGEEDDLATRTSMAG-LKVSRYP 304
DB 637 GFDTKIEGMGKEVDVDFEKAIRKGRRLVIRVP 668

RESULT 12

H82884
multiple banded antigen homolog UR483 [imported] - Ureaplasma urealyticum
C:Species: Ureaplasma urealyticum
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: H82884
R:Glaser, J.L.; Lefkowitz, E.J.; Glaser, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A:Reference number: A82870
A:Accession: H82884
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1103 <GLA>
A:Cross-references: UNIPARC:UPI00000C1C8A; GB:AE02146; GB:AF222894; NID:G6899479; PIDD:
A:Experimental source: serovar 3; biovar 1
C:Genetic:
A:Gene: UR483
A:Genetic code: SGC3

Query Match 5.1%; Score 103; DB 2; Length 1103;

Best Local Similarity 22.7%; Pred. No. 2.1;

Matches 82; Conservative 54; Mismatches 104; Indels 122; Gaps 23;

QY 71 NSTFSPIS-EVNGSFEMEDIRPLFPNQLTFQNGPRPHVGRIRVFLDEDPKTELEKI 129
DB 739 NQTFKLISFEIDDNVY--TNNTVKTILNESSPHLL--LNLEFNDIKNTNST 786
QY 130 YPPT-HAGGGMPCDVCARRVALIVP-----YDREAH-LRIMLHLH- 171
DB 787 LKLTINQAVINMR-----RVDVILKKNKTKNKLILFKPDYFDKXTHQISTGLDLNLI 840
QY 172 -----SLAKQOLDYAIIFYEQVANQTFNRGKLMNVGYD-----V 206
DB 841 NTDYEISELFIDKQIN-----LKNINNTKF--KTLDYGPDPYLDKIDHEINQOQKNI 892
QY 207 ASRLYPWQCFI-----HDVDLPRDRLNLTGPIQPRHMSVAD-----KFNLYLPSAI 257
DB 893 KFKLH-FSDIFKDKNAHKSLVLEDSNQQECIDLNAHQKLDIYEHQTIPEFLKY-- 948

QY 258 FCGISALTRDHLKINGFS-NDPFWGGEEDDLATRTSMAGLKVSRYPTQIARYKMKHS 316
DB 949 -----LKRDKYRINSLQAND-----KNULD--TWVDNQRYNFPQIL----- 984
QY 317 TEATNP-----VNKCRYKIMG-OTGRWRTRDG---LSNLKY-KLVNLEKPLIYTRAVNVL 366
DB 985 ---NPNLWKVNDNIKYEYVNGSDPKINITFDODIFLSNLSODKECNIELNLYVYNNSDGT 1040
QY 367 LE 368
DB 1041 LE 1042

RESULT 13

T42245
probable polypeptide N-acetylglucosaminyltransferase (EC 2.4.1.41) - Caenorhabditis ele
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42245
R:Hagen, F.K.; Nehrke, K.
J: Biol. Chem. 273, 8268-8277, 1998
A:Title: cDNA cloning and expression of a family of UDP-N-acetyl-Dgalactosamine:Polypecti
A:Reference number: Z22126; MUID:98192620; PMID:9525933
A:Accession: T42245
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-623 <HAG>
A:Cross-references: UNIPROT:061391; UNIPARC:UPI000008050F; EMBL:AF031835; NID:G3047190; I
C:Genetic:
A:Gene: gly-5
C:Superfamily: polypeptide N-acetylglucosaminyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 4.9%; Score 99.5; DB 2; Length 623;

Best Local Similarity 22.2%; Pred. No. 1.9; 107; Indels 85; Gaps 18;

Matches 69; Conservative 50; Mismatches 95;

QY 105 QTFPHVGPRIYFLDEPDP---KTLEKIYPTTHAG-----HGMPKDCVARRHY 150
DB 201 RTRDHLLEEVLVDDSDMDHTKRPLEBYM--SQFGKVKILMEKREGILR---ARLKG 255
QY 151 AI-----IYPRDREAHRLIMLHLSLAKQOLDYAI--IVEQVANQTF---NRKG 198
DB 256 AAVATGEVLTLYLD--SHCEMGEEMERLDRIRKRDPTTVCPIVDVINDTEFHNHNSKAY 313
QY 199 LMNVGYDVASRLYPWQC-FIPHDVDDLPRDRLNLTGPIQPRHMSVADIKFNLYLPSAI 257
DB 314 FTSVGG-----FDWGLQFNWHSI--PERDRKNRTRPIDP-----VRSPTMA- 352
QY 258 FCGISALTRDHLKINGFSNDPFWGGEEDDLATRTSMAGLKVSRYPTQIARYKMKHST 317
DB 353 -GGLSFIDKYEPEKLTYPDGFIDWGENLELSFKLMCGGLEIYVCSHGVHFFKRSP 411
QY 318 EATNPVNCKRYKMGOTKRRWTRDGLSNLKYYKLVN-----ELKLYTRAVVDL----- 367
DB 412 -----YK-----W-RTGVNVYLRKNSIRLAEWLDDYKYYERINNQLGDFG 452
QY 368 EKDCRELRND 378
DB 453 DISSRKRLRND 463

RESULT 14

T42247
polypeptide N-acetylglucosaminyltransferase (EC 2.4.1.41) - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42247
R:Hagen, F.K.; Nehrke, K.
J: Biol. Chem. 273, 8268-8277, 1998
A:Title: cDNA cloning and expression of a family of UDP-N-acetyl-Dgalactosamine:Polypecti
A:Reference number: Z22126; MUID:98192620; PMID:9525933
A:Accession: T42247

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-624 <HAG>
A:Cross-references: UNIPROT:O61393; UNIPARC:UPI000007D64; EMBL:AF031837; NID:G3047194;
C:Genetics:
A:Gene: gly-5
C:Superfamily: polypeptide N-acetylglucosaminyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 4.9%; Score 99.5; DB 2; Length 624;
Best Local Similarity 22.2%; Pred. No. 1.9;
Matches 69; Conservative 50; Mismatches 107; Indels 85; Gaps 18;

QY 105 QTPPHLVGPIRVFLDEPDF---KTLEKITYPTTHAGS-----HGMPKDCVARRHV 150
DB 201 RTPDHLLEEVVLVDPSMDHTKRPLEETM--SQFGKVKILMEKREGLIR---AKLRG 255
QY 151 AI-----IVPYDRRAHLRIMLNHLSLAKOQDYAIF--IVEQVANQTF---NRGK 198
DB 256 AAVATGEVLTYLD--SHCECMGMEPILDRIKRDPPTVCPVIDIDNTFEYHHSKAY 313
QY 199 LMMNGDVASRLYPWQC-FIFHDVLLPEDDRLTYCPIQPRHMSVAIDKFNKLPYSAI 257
DB 314 FTSVGG-----FDWGLQFNWHSI---PRDRKNRTRPIDP-----VRSEPTMA- 352
QY 258 FGGISALTKDHLKKINGFNDPMWGGGEDDLATRTSMAGLKVSRYPTQIARYKMIKHST 317
DB 353 -GGLSIDKEYPEKLGTYDPGFDIWGENLEISFKIMCGGTLEIVPCSHGVHFRKRSF 411
QY 318 EATNPVNKCRYKIMGQTKRRWTRDGLSNLKYYLVN-----ELKPLYTRAVVDLL--- 367
DB 412 -----YK-----W-RTGVNVLKNSIRLAEVWLDYKTYYYERINNQLGDRG 452
QY 368 EKDCRRELARD 378
DB 453 DISSRKCLRED 463

RESULT 15

T42246
Polypeptide N-acetylglucosaminyltransferase (EC 2.4.1.41) - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42246
R:Hagen, F.K.; Nehrke, K.
J: Biol. Chem. 273, 8268-8277, 1998
A:Title: cDNA cloning and expression of a family of UDP-N-acetyl-D-glucosamine:Polypeptide
A:Reference number: Z22126; MUID:98192620; PMID:9525933
A:Accession: T42246
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-626 <HAG>
A:Cross-references: UNIPROT:O61392; UNIPARC:UPI0000078C85; EMBL:AF031836; NID:G3047192;
C:Genetics:
A:Gene: gly-5
C:Superfamily: polypeptide N-acetylglucosaminyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 4.9%; Score 99.5; DB 2; Length 626;
Best Local Similarity 22.2%; Pred. No. 1.9;
Matches 69; Conservative 50; Mismatches 107; Indels 85; Gaps 18;

QY 105 QTPPHLVGPIRVFLDEPDF---KTLEKITYPTTHAGS-----HGMPKDCVARRHV 150
DB 201 RTPDHLLEEVVLVDPSMDHTKRPLEETM--SQFGKVKILMEKREGLIR---AKLRG 255
QY 151 AI-----IVPYDRRAHLRIMLNHLSLAKOQDYAIF--IVEQVANQTF---NRGK 198
DB 256 AAVATGEVLTYLD--SHCECMGMEPILDRIKRDPPTVCPVIDIDNTFEYHHSKAY 313
QY 199 LMMNGDVASRLYPWQC-FIFHDVLLPEDDRLTYCPIQPRHMSVAIDKFNKLPYSAI 257
DB 314 FTSVGG-----FDWGLQFNWHSI---PRDRKNRTRPIDP-----VRSEPTMA- 352

QY 258 FGGISALTKDHLKKINGFNDPMWGGGEDDLATRTSMAGLKVSRYPTQIARYKMIKHST 317
DB 353 -GGLSIDKEYPEKLGTYDPGFDIWGENLEISFKIMCGGTLEIVPCSHGVHFRKRSF 411
QY 318 EATNPVNKCRYKIMGQTKRRWTRDGLSNLKYYLVN-----ELKPLYTRAVVDLL--- 367
DB 412 -----YK-----W-RTGVNVLKNSIRLAEVWLDYKTYYYERINNQLGDRG 452
QY 368 EKDCRRELARD 378
DB 453 DISSRKCLRED 463

Search completed: January 6, 2006, 15:23:14
Job time : 42 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:11:03 ; Search time 133 Seconds
(without alignments)
1265.279 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSILVLCAY.....VDLKKCRRLRPPTCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_21.*
1: geneseqp1980s:*
2: geneseqp1908s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	383	8	ADM32968
2	718	35.4	404	3	AA843627
3	717	35.3	393	8	ABM80189
4	710	35.0	393	8	ABE85868
5	709	34.9	403	4	AB863921
6	680.5	33.5	362	7	ADG63356
7	680.5	33.5	362	7	ADL27179
8	673.5	33.2	362	9	ABE95869
9	668.5	32.9	264	3	ABE95869
10	668.5	32.9	767	2	AA855709
11	668	32.9	398	2	AA855709
12	667.5	32.9	352	7	ABM85878
13	667.5	32.9	399	8	ADQ79982
14	667.5	32.9	399	8	ADQ79982
15	667.5	32.9	399	8	AEAS5054
16	667.5	32.9	399	8	AEAS5054
17	666	32.8	385	3	AB803647
18	666	32.8	385	3	AB803647
19	666	32.8	398	7	AA828838
20	666	32.8	398	7	ADP18448
21	666	32.8	398	7	ADP18448
22	666	32.8	398	8	ADQ78358
23	666	32.8	398	8	ADQ78358
24	665.5	32.8	767	2	AA855708

25	665	32.7	398	2	AA855706	AA855706	Galactose
26	663.5	32.7	371	8	ADQ91374	ADQ91374	Manit-Gal
27	661	32.5	475	5	ADP18451	ADP18451	N-acetyl
28	661	32.5	500	7	ADP18450	ADP18450	N-acetyl
29	660.5	32.5	365	4	AAU01864	AAU01864	MAN_ATG
30	660.5	32.5	397	9	ADZ69855	ADZ69855	Human Gal
31	660.5	32.5	402	4	AAU01865	AAU01865	MNS285 bo
32	660.5	32.5	406	4	AAU01867	AAU01867	GNT galac
33	660.5	32.5	433	4	AAU01866	AAU01866	MAN GMGT
34	660	32.5	372	7	ADJ94854	ADJ94854	NOV
35	657	32.3	372	6	ABU08239	ABU08239	Human Bet
36	657	32.3	372	7	AAE39165	AAE39165	Human Bet
37	657	32.3	372	8	ABM80077	ABM80077	Tumour-as
38	657	32.3	372	8	ABE95865	ABE95865	Human bet
39	656.5	32.3	720	8	ADQ78360	ADQ78360	alpha-1,3
40	655.5	32.3	355	7	ADD68135	ADD68135	Human gal
41	655.5	32.3	383	7	ADD68087	ADD68087	Human gal
42	655.5	32.3	383	7	ADD68141	ADD68141	Human gal
43	655	32.3	398	8	ADQ79981	ADQ79981	Human bet
44	655	32.3	398	8	AEAS1683	AEAS1683	Human bet
45	655	32.3	402	2	AA805933	AA805933	Mouse bet

ALIGNMENTS

RESULT 1	ADM32968	ADM32968	standard, protein, 383 AA.
ID	ADM32968	ADM32968	standard, protein, 383 AA.
XX	ADM32968	ADM32968	standard, protein, 383 AA.
AC	ADM32968	ADM32968	standard, protein, 383 AA.
XX	ADM32968	ADM32968	standard, protein, 383 AA.
DT	17-JUN-2004	(first entry)	
XX	17-JUN-2004	(first entry)	
DE	Amino acid sequence of a beta1,4-N-acetylglucosaminyltransferase.		
XX	beta1, 4-N-acetylglucosaminyltransferase; beta4GalNAcT; animal cell;		
KW	glycoprotein; lactinac motif; IDN motif; GalNAc-beta4GalNAc-R;		
KW	glycoprotein; hormone; pituitary gland; glycoconjugate; parasite;		
KW	immune response; infection; enzyme.		
XX	Caenorhabditis elegans.		
OS	Caenorhabditis elegans.		
XX	Caenorhabditis elegans.		
PH	Key	Location/Qualifiers	
FT	Domain	7..29	
FT	Domain	/note = putative transmembrane domain	
FT	Modified-site	37	
FT	Modified-site	/note= "N-glycosylation site"	
FT	Region	48..50	
FT	Region	/note= "DVD motif"	
FT	Modified-site	71	
FT	Modified-site	/note= "N-glycosylation site"	
FT	Modified-site	81	
FT	Modified-site	/note= "N-glycosylation site"	
FT	Modified-site	97	
FT	Modified-site	/note= "N-glycosylation site"	
FT	Modified-site	104	
FT	Modified-site	/note= "N-glycosylation site"	
FT	Modified-site	191	
FT	Modified-site	/note= "N-glycosylation site"	
FT	Region	220..222	
FT	Region	/note= "DVD motif"	
PN	WO2004024938-A2.		
XX	WO2004024938-A2.		
PD	25-MAR-2004.		
XX	25-MAR-2004.		
XX	12-SEP-2003; 2003WO-US028833.		
PF	12-SEP-2003; 2003WO-US028833.		
XX	13-SEP-2002; 2002US-0411242P.		
XX	13-SEP-2002; 2002US-0411242P.		
XX	(CDMM/) CUMMINGS R D.		
XX	(CDMM/) CUMMINGS R D.		
PA	(KAWA/) KAWAR Z.		
PA	(KAWA/) KAWAR Z.		

XX Cummings RD, Kavar Z;
XX
XX WPI, 2004-270055/25.
XX N-PSDB; ADM32969.
XX
PT New purified approximatelyb4 acetylglucosamyl transferase that is
PT substantially free of other protein, useful for transforming or
PT transfecting host cells for producing substantially pure forms of the
PT enzyme.
XX
XX Claim 2; Fig 1; s1pp; English.
XX
XX The present sequence represents a beta1,4-N-
XX acetylglucosaminyltransferase, designated beta4GALNAcT. The enzyme is
XX required for the biosynthesis of animal cell glycoproteins. The enzyme
XX functions to synthesise the lactinac or LCN motif GalNAcbeta4GlcNAc-R.
XX This motif is glycoprotein hormones produced by the pituitary gland, and
XX is a major marker of glycoconjugates made by parasitic and non-parasitic
XX invertebrates and may be implicated in host immune responses to
XX infection. The protein and polynucleotides are useful for transforming or
XX transfecting host cells for producing substantially pure forms of the
XX enzyme, or in vitro, for formation of a LCN structure on proteins or
XX peptides.
XX
XX Sequence 383 AA:
SQ
Query Match 100.0%; Score 2031; DB 8; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.4e-208;
Matches 383; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAFRHLAVARLKSLLVCAVLLVHAMITKPSITENLTIGSSSTLADVAMEAVLGNTA 60
DB 1 MAFRHLAVARLKSLLVCAVLLVHAMITKPSITENLTIGSSSTLADVAMEAVLGNTA 60
QY 61 STSDDLDTWNSSTFSPISSEVNGTSFMEDIRPILFPDNOQLQFCNQTPRHVGPRIVFLDE 120
DB 61 STSDDLDTWNSSTFSPISSEVNGTSFMEDIRPILFPDNOQLQFCNQTPRHVGPRIVFLDE 120
QY 121 PDKRTLEKTPDTHAGHGMPKOCVAHRVAIIVPYDRRAHLRIMAHNLHSLAAQQLD 180
DB 121 PDKRTLEKTPDTHAGHGMPKOCVAHRVAIIVPYDRRAHLRIMAHNLHSLAAQQLD 180
QY 181 YAFIVEQVANQTFNRKGLANVGYDVASRLVPMQCFIFHVDLLPEDDRNLVTCPIOPRH 240
DB 181 YAFIVEQVANQTFNRKGLANVGYDVASRLVPMQCFIFHVDLLPEDDRNLVTCPIOPRH 240
QY 241 MSVAIDKFNKPLPYSAIFGGISALTQHLKKKINGFSNDPFGWGGEEDDLATRTSMAGLKV 300
DB 241 MSVAIDKFNKPLPYSAIFGGISALTQHLKKKINGFSNDPFGWGGEEDDLATRTSMAGLKV 300
QY 301 SRPYTOIARFKMTKHSREATNPVKKCRKIMGOTKRRWTDGLSNLKKLVNELKRLYT 360
DB 301 SRPYTOIARFKMTKHSREATNPVKKCRKIMGOTKRRWTDGLSNLKKLVNELKRLYT 360
QY 361 RAVVDLLEKDCRRELKRDPTCF 383
DB 361 RAVVDLLEKDCRRELKRDPTCF 383
RESULT 2
AAB43627
ID AAB43627 standard; protein, 404 AA.
XX
XX AAB43627;
XX
XX 08-FEB-2001 (first entry)
DE Human cancer associated protein sequence SEQ ID NO:1072.
XX
XX Human, cancer associated gene; cancer antigen, detection; cancer;
XX diagnosis; cytotoxic; proliferative; vulnerrary; immunomodulator;
XX antidiabetic; antiaesthetic; antirheumatic; antiarthritic; antiviral;
XX

KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening.
XX
XX Homo sapiens.
XX
XX WO20005350-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US005882.
XX
XX 12-MAR-1999; 99US-0124270P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI, 2000-587533/55.
XX N-PSDB; AAC77836.
XX
PT Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
XX
XX Claim 11; Page 1670-1671, 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
XX AAB43398 to AAB44239. The proteins can have activities based on the
XX tissues and cells the genes are expressed in. Example of activities
XX include: cytostatic; proliferative; vulnerrary; immunomodulator;
XX antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
XX vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
XX polynucleotides and polypeptides can be used for preventing, treating or
XX ameliorating medical conditions and diagnosing pathological conditions.
XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from
XX the present invention may be used to treat immune disorders by activating
XX or inhibiting the proliferation, differentiation or mobilisation of
XX immune cells, to treat disorders of haematopoietic cells, autoimmune
XX disorders, allergic reactions, graft versus host disease and organ
XX rejection, modulate haemostatic or thrombolytic activity, modulate
XX inflammation, cancer, cardiovascular disorders, neurological disease and
XX bacterial or viral infections. The peptides, nucleotides, antibodies,
XX agonists and antagonists may be also be used in drug screens. AAC78449 to
XX the present invention
XX
XX Sequence 404 AA:
SQ
Query Match 35.4%; Score 718; DB 3; Length 404;
Best Local Similarity 48.9%; Pred. No. 1.7e-67;
Matches 134; Conservative 43; Mismatches 95; Indels 2; Gaps 2;
QY 95 PDNOQLQFCNQTPRHVGPRIV-FLDEPDKTLEKTPDTHAGHGMPKOCVAHRVAIT 153
DB 80 FXPDGLPYCERSPLLVGPVSFSPVPSIAETVENNPVPEGRYRPPGCEPERSRTAIL 139
QY 154 VPYRDEARHLRIMAHNLHSLAAQQLDYAFIVEQVANQTFNRKGLANVGYDVASRLVPM 213
DB 140 VPHRAREHRLRLYLHLPLOQQLAYGYVVIHQNGNGFNPAKLLNNGVRAALDEEW 199
QY 214 QCFTPHDVDLLPEDDRNLVTC-PIQPHMSVAIDKFNKPLPYSAIFGGISALTQHLKKT 272
DB 200 DCLFLHDVDLLPENDDNLVYCDPRGPRHVAVANMKFGYSLLPYQYGGYSALTTPDYLYKM 259
QY 273 NGFSNPFMGWGGEEDDLATRTSMAGLKVSRYPQIARFKMTKHSREATNPVKKCRKIMG 332
DB 260 NGPNEYMGWGGEEDDLATRTSMAGLKVSRYPQIARFKMTKHSREATNPVKKCRKIMG 319

QY 90 RPLFPDNGTLQFCNQTPPHLVGPIRVFLDER-DEKTELEKIYPDTHAGHGMPKDCVARN 148
 Db 80 GPRNLPDSAPSGLCDFPSPFLVGLVPLKVERSGQVNLBEVASTNPEVREGRRFAPKCKALQ 139
 QY 149 RYAIIVPYRDRBAHLRIMLHNLHSLAKQOLDYAFIVEQVANTFNKGLMANGVDVAS 208
 Db 140 KVAIIIPFRNREBHKYWLVMHPILOQOLDYGVVAINQDDBEFNRKALINVGFTAL 199
 QY 209 RLYPMQCFIFHDVDLLPEDDRNLVTCPIQPRHMSVAIDKPNYKLPYSALFGISALTCDH 268
 Db 200 KEYDYDCEVFSVDVLLPMDDBRTYKCYSPRHLVSMDKFGRLPNQYFGVSALSKEQ 259
 QY 269 LKINGFNSDFWGWGEGDDDLATRTSMAGIKVSRYPQTQARXKMTKHSYEAATNPVKKRY 328
 Db 260 FTKINGFPNNYMGWGEDDDIYNRLVFKMGISRPDAVIGKCRMLRHSRDRNENPERF 319
 QY 329 KIMGQTKRRWTDGSLNKKYKLVNLEKPLVTRAVVDL 366
 Db 330 DRIATHRTMSSDGLNSLSYEVLRDTRFPLVTRITVDI 357

RESULT 7
 ID ADL27179 standard; protein; 362 AA.
 AC ADL27179;
 XX 20-MAY-2004 (first entry)
 DT 20-MAY-2004 (first entry)
 XX Chicken beta1,4-galactosyltransferase.
 DE Chicken beta1,4-galactosyltransferase.
 XX chicken; galactosyltransferase; beta1,4-galactosyltransferase; cancer;
 KM adenocarcinoma; leukemia; lymphoma; melanoma; developmental disorder;
 KM renal tubular acidosis; anaemia; Cushing's syndrome; Wilms' tumour;
 KM aniridia; mental retardation syndrome; congenital glaucoma; cataract;
 KM reproductive disorder; infertility; endometriosis;
 KM polycystic ovary syndrome; uterine fibroid; Peyronie's disease;
 KM benign prostatic hyperplasia; autoimmune disorder; inflammatory disorder;
 KM AIDS; Addison's disease; asthma; atherosclerosis; bronchitis;
 KM Crohn's disease; atopic dermatitis; Grave's disease; multiple sclerosis;
 KM myasthenia gravis; Werner syndrome.
 KM Gallus gallus.
 OS Gallus gallus.
 XX US2003203396-A1.
 PN 30-OCT-2003.
 PD 30-OCT-2003.
 XX 10-JUN-2003; 2003US-00459311.
 PF 03-APR-1998; 98US-00055097.
 PR 12-AUG-1999; 99US-00373902.
 XX (INCY-) INCYTE CORP.
 PA Hillman JL, Guegler KJ, Corley NC, Shah P, Patterson C;
 PI WPI; 2003-900634/82.
 DR Novel human galactosyltransferase useful for treating or preventing
 XX developmental disorders such as anemia, mental retardation, cancer such
 XX as leukemia, lymphoma, reproductive disorders, autoimmune disorders.
 XX Disclosure; SEQ ID NO 7; 49pp; English.
 PS The invention relates to an isolated human galactosyltransferase
 XX polypeptide (I). (I) is useful for screening a compound for effectiveness
 XX as an agonist or antagonist of (I). (I) is useful for screening a
 XX compound that specifically binds to (I). (I) is useful for screening for
 XX a compound that modulates the activity of (I). (I) is useful for
 XX preparing polyclonal antibody (V). (V) is useful for a diagnostic test
 XX for a condition or disease associated with the expression of human
 XX galactosyltransferase (HUGA) in a biological sample. (V) is useful for

CC detecting (I) in sample and purifying (I) from a sample. (I) or its
 CC fragment or derivative is useful for treating cancer such as
 CC adenocarcinoma, leukemia, lymphoma, melanoma, developmental disorders
 CC such as renal tubular acidosis, anaemia, Cushing's syndrome, Wilms'
 CC tumour, aniridia, mental retardation syndrome, congenital glaucoma,
 CC cataract, reproductive disorders such as infertility, endometriosis,
 CC dysplasia of uterous cycle, polycystic ovary syndrome, uterine
 CC fibroids, Peyronie's disease, benign prostatic hyperplasia and
 CC autoimmune/inflammatory disorders such as AIDS, Addison's disease,
 CC asthma, atherosclerosis, bronchitis, Crohn's disease, atopic dermatitis,
 CC Grave's disease, multiple sclerosis, myasthenia gravis, Werner syndrome.
 CC The present sequence represents the amino acid sequence of Chicken
 CC beta1,4-galactosyltransferase.
 XX Sequence 362 AA;
 SQ

Query Match 33.5%; Score 680.5; DB 7; Length 362;
 Best Local Similarity 43.9%; Pred. No. 1.5e-63;
 Matches 122; Conservative 58; Mismatches 97; Indels 1; Gaps 1;

QY 90 RPLFPDNGTLQFCNQTPPHLVGPIRVFLDER-DEKTELEKIYPDTHAGHGMPKDCVARN 148
 Db 80 GPRNLPDSAPSGLCDFPSPFLVGLVPLKVERSGQVNLBEVASTNPEVREGRRFAPKCKALQ 139
 QY 149 RYAIIVPYRDRBAHLRIMLHNLHSLAKQOLDYAFIVEQVANTFNKGLMANGVDVAS 208
 Db 140 KVAIIIPFRNREBHKYWLVMHPILOQOLDYGVVAINQDDBEFNRKALINVGFTAL 199
 QY 209 RLYPMQCFIFHDVDLLPEDDRNLVTCPIQPRHMSVAIDKPNYKLPYSALFGISALTCDH 268
 Db 200 KEYDYDCEVFSVDVLLPMDDBRTYKCYSPRHLVSMDKFGRLPNQYFGVSALSKEQ 259
 QY 269 LKINGFNSDFWGWGEGDDDLATRTSMAGIKVSRYPQTQARXKMTKHSYEAATNPVKKRY 328
 Db 260 FTKINGFPNNYMGWGEDDDIYNRLVFKMGISRPDAVIGKCRMLRHSRDRNENPERF 319
 QY 329 KIMGQTKRRWTDGSLNKKYKLVNLEKPLVTRAVVDL 366
 Db 330 DRIATHRTMSSDGLNSLSYEVLRDTRFPLVTRITVDI 357

RESULT 8
 ID AEB95869 standard; protein; 362 AA.
 AC AEB95869;
 XX 20-OCT-2005 (first entry)
 DT 20-OCT-2005 (first entry)
 XX Chicken beta-1,4-galactosyltransferase (CKI) protein, SEQ ID NO: 6.
 DE Chicken beta-1,4-galactosyltransferase (CKI) protein, SEQ ID NO: 6.
 XX DNA polymorphism; enzyme; beta-1,4-galactosyltransferase.
 KM Gallus gallus.
 OS Gallus gallus.
 XX US2005181437-A1.
 PN 18-AUG-2005.
 PD 13-APR-2005; 2005US-00105796.
 PF 17-JUL-1998; 98US-00118464.
 PR 24-APR-2002; 2002US-00132652.
 XX (GLYC-) GLYCOTYM APS.
 PA Clausen H, Bennett EP;
 PI WPI; 2005-563212/57.
 DR GENBANK; AAB05218.
 XX New nucleic acid encoding UDP-galactose: beta-N-acetylglucosamine beta-
 XX 1,4-galactosyltransferase, useful for identifying DNA polymorphism in

PR 27-NOV-1992; 92EP-00810924.
XX (CIBA) CIBA GEIGY AG.
XX Berger EG, Matzele M, Iwanow SX;
XX WPI, 1994-200274/24.
DR N-PSDB; AAO66892.
XX
XX Proteins with glycosyl transferase activity - useful for synthesis or
PT modification of glyco-proteins, glyco-lipid(s) and oligosaccharide(s).
PS Disclosure; Page 57-61; 67pp; English.
XX
XX Hybrid glycosyltransferases (see also AARS5708) consisting of a membrane-
CC bound galactosyltransferase linked at its C-terminal to the N-terminal of
CC a soluble sialyltransferase have been expressed in Saccharomyces
CC cerevisiae using plasmid YEPGSta. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 767 AA;
Query Match 32.9%; Score 666.5; DB 2; Length 767;
Best Local Similarity 35.4%; Pred. No. 9,5e-62;
Matches 140; Conservative 74; Mismatches 132; Indels 49; Gaps 6;
QY 14 LLVCAVLLVHAMTY-----KIPSLYENLTIGSSSTLIADVMEAVLGNTA----- 60
DB 26 LVAVCALHLGVTLYVYLAGRDISRLPOL-----VGVSPTLQGSNSAAAIQSSGSELRTG 80
QY 61 -----STSDLLDTMNSTFSPISSEVNOTSFMEDIRPILFPDNOTLQFC 103
DB 81 GARPPPLGASSQPRRGDSSPVVDSGPGPASNLTSV-----PVPHTTALSIPAC 130
QY 104 NOTPPLHVGPIRVFLDEP-DFKTELEKIYPTDHAGHGMPKDCVARRVAIIIVPYRDREAH 162
DB 131 PEESPLVGPMLIEFMNPVDLELVAKONPNVMGGRVAPRDCVSPHKVAIIIPFNROGH 190
QY 163 LRIMAHNSLAKQOOLDAIFIVGVANOTFNRGKLMNVGDVASRLYPMOCFFHYVD 222
DB 191 LKTYWLYYHLPVLRQQLDYGIVINQAGDTIFNRKALNVGQELAKDYDTCFVFSVD 250
QY 223 LLPEDDRLNYTCPIOPRHMSVAIDKFNKYLPYSAIFGGSALTQKHLKKINGFSDFMGM 282
DB 251 LIPMDHNAYRCFSQPRHISVAMDKFGSLPYVQYFGVSALSQKQPLTINGFPNNYMGW 310
QY 283 GGEEDDLATRTSMAGLKVSRYPTQIARYKMIKHSSTEATNPVNCRYKIMGQTKRRWTRDG 342
DB 311 GGEEDDIFNRLVFRGMSISRPNAAVGRCRMIRHSRDKNEPQRPDRIAHTKETMLSDG 370
QY 343 LSNLKYKLVNLEKPLTYRAVVDLEKCRRELRR 377
DB 371 LNSTLYQVLDVQRYPLTYQITVDI---GTRAIRR 402
RESULT 11
AA17862
ID AA17862 standard; protein; 398 AA.
XX
XX AA17862;
AC
XX 17-AUG-1999 (first entry)
DT
XX Human beta-1,4-galactose transferase.
DE
XX Human; beta-1,4-galactose transferase; preparation.
KM
XX Homo sapiens.
OS
XX JP11137247-A.
PN
XX 25-MAY-1999.
PD
XX

PF 10-NOV-1997; 97JP-00306967.
XX
XX 10-NOV-1997; 97JP-00306967.
XX
XX (TOYM) TOYOKO KK.
XX WPI, 1999-374371/32.
DR N-PSDB; AAX80151.
XX
XX Preparing beta-1,4-galactose transferase - using recombinant techniques.
PT
XX
XX Disclosure; Page 7-8; 9pp; Japanese.
XX
XX A method has been developed for the preparation of human-derived beta-1,4-
CC galactose transferase. The method comprises transformation of
CC Escherichia coli by an expression vector containing a gene encoding human
CC -derived beta-1,4-galactose transferase and a gene coding maltose-
CC combined protein. The transformant is cultured to form a fusion protein
CC consisting of human-derived beta-1,4-galactose transferase and maltose-
CC combined protein. The fusion protein is purified by affinity
CC chromatography and digested with enzymes to form beta-1,4-galactose
CC transferase. The method can be used to prepare human-derived beta-1,4-
CC galactose transferase easily and efficiently in large amounts. The
CC present sequence represents human beta-1,4-galactose transferase
SQ Sequence 398 AA;
Query Match 32.9%; Score 668; DB 2; Length 398;
Best Local Similarity 35.7%; Pred. No. 3,9e-62;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;
QY 14 LLVCAVLLVHAMTY-----KIPSLYENLTIGSSSTLIADVMEAVLGNTA----- 60
DB 26 LVAVCALHLRVTLVYVYLAGRDISRLPOL-----VGVSPTLQGSNSAAAIQSSGSELRTG 80
QY 61 -----STSDLLDTMNSTFSPISSEVNOTSFMEDIRPILFPDNOTLQFC 103
DB 81 GARPPPLGASSQPRRGDSSPVVDSGPGPASNLTSV-----PVPHTTALSIPAC 130
QY 104 NOTPPLHVGPIRVFLDEP-DFKTELEKIYPTDHAGHGMPKDCVARRVAIIIVPYRDREAH 162
DB 131 PEESPLVGPMLIEFMNPVDLELVAKONPNVMGGRVAPRDCVSPHKVAIIIPFNROGH 190
QY 163 LRIMAHNSLAKQOOLDAIFIVGVANOTFNRGKLMNVGDVASRLYPMOCFFHYVD 222
DB 191 LKTYWLYYHLPVLRQQLDYGIVINQAGDTIFNRKALNVGQELAKDYDTCFVFSVD 250
QY 223 LLPEDDRLNYTCPIOPRHMSVAIDKFNKYLPYSAIFGGSALTQKHLKKINGFSDFMGM 282
DB 251 LIPMDHNAYRCFSQPRHISVAMDKFGSLPYVQYFGVSALSQKQPLTINGFPNNYMGW 310
QY 283 GGEEDDLATRTSMAGLKVSRYPTQIARYKMIKHSSTEATNPVNCRYKIMGQTKRRWTRDG 342
DB 311 GGEEDDIFNRLVFRGMSISRPNAAVGRCRMIRHSRDKNEPQRPDRIAHTKETMLSDG 370
QY 343 LSNLKYKLVNLEKPLTYRAVVDL 366
DB 371 LNSTLYQVLDVQRYPLTYQITVDI 394
RESULT 12
ABM85878
ID ABM85878 standard; protein; 352 AA.
XX
XX ABM85878;
AC
XX 18-NOV-2004 (first entry)
DT
XX Mouse protein sequence mCp12786.
DE
XX Cytostatic; carcinoma; lymphoma; cancer; murine.
KM
XX
XX Mus musculus.
OS

Db	367	MRFDGLNSLFYKVLVDVQRVPLTYQTITVDI	395
		RESULT 14	
ID	AEAS5054	AEAS5054 standard; protein; 399 AA.	
XX	AEAS5054;		
XX	11-AUG-2005	(first entry)	
XX	Mouse beta-1, 4-galactosyl transferase 1 protein, SEQ ID NO: 30.		
XX	plasma membrane; diagnosis; therapeutic; cancer; cytostatic; neoplasm;		
XX	beta-1, 4-galactosyl transferase 1; enzyme.		
XX	Mus musculus.		
XX	MO2005052182-A2.		
XX	09-JUN-2005.		
XX	25-NOV-2004; 2004MO-IL001085.		
XX	26-NOV-2003; 2003US-052486SP.		
XX	(YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.		
XX	Linfal M, Imberg A, Bledt Y;		
XX	WPI; 2005-418017/42.		
XX	SWISSPROT; P15535.		
XX	Characterizing proteins present in a plasma membrane of a cell, useful in		
XX	identifying diagnostic markers and potential drugs, comprises subjecting		
XX	a cell to a protease treatment.		
XX	Claim 25; SEQ ID NO 30; 196bp; English.		
XX	The present invention relates to a method of characterizing proteins		
XX	present in the plasma membrane (PM) of live cells. The proteins of the		
XX	invention are useful in identifying diagnostic markers and potential		
XX	drugs. The invention is useful for identifying drugs for diagnosing and		
XX	treating disorders such as cancer which are associated with abnormal		
XX	representation of cell surface proteins. The present sequence is mouse		
XX	beta-1, 4-galactosyl transferase 1 (also referred as beta-1, 4-GalTase)		
XX	protein.		
XX	Sequence 399 AA;		
XX	Query Match	32.9%; Score 667.5; DB 9; Length 399;	
XX	Best Local Similarity	45.7%; Pred. No. 4,4e-62;	
XX	Matches 123; Conservative 52; Mismatches 93; Indels 1; Gaps 1		
Qy	99 TLQFCNQTTPHVLGVPIRV-FLDEPDPFKTLKITYPTTHAGHGMPKDCVARRHVAIIIVPYR	157	
Db	127 SLPACPEBSPPLVCGMILIDFNIAVDLELLAKKNPEIKTGGRYSPKDCVSPHKVAILIIPFR	186	
Qy	158 DREAIRLRIMNTHSLAKQOUDYALFIVEQVANOFTPRGKLMTNGYDVASLTYMOCFI	217	
Db	187 NRQELHAKWLYLHPIILORQOUDYGIYVINOAGDTMFPRAKLNLGFOELAKLDYVYNCFV	246	
Qy	218 FHDVDLLPEDDRNLTCTCPILQPRHMSVALIDKFNYKLPYSALFGGISALTGKHLLKINGFSN	277	
Db	247 FSDVDLLIMDDRNARVRCFSQPRHISVANDKFGSLPTVQYFGGVASLSKQFLAINGFPN	306	
Qy	278 DFWMGSGEDDDLATFTSMAGLKVSSVPTQIARYYKIKISTEATNPNVNNCKRYKIMQOTRR	337	
Db	307 NYWGGSGEDDDIFNLTVHKGMSISPRNAVVGCRMIIRSRDKKNPNPQRFRIAHFTKET	366	
Qy	338 WTRDGLSNLKYKLVNLVLELKLPTTRAVVDL	366	
Db	367 MRFDGLNSLFYKVLVDVQRVPLTYQTITVDI	395	

RESULT 15	
AEAS1684	..
ID	AEAS1684 standard; protein; 399 AA.
XX	
AC	AEAS1684;
XX	
DT	25-AUG-2005 (first entry)
XX	
DE	Mouse beta(1,4)-galactosyltransferase I.
XX	
KM	protein purification; vaccine; immunity; immunogenicity;
XX	beta(1,4)-galactosyltransferase I; enzyme.
XX	
OS	Mus musculus.
XX	
PH	Location/Qualifiers
FT	Misc-difference 225
FT	/note="Arg of the catalytic domain can be exchanged for
FT	another amino acid residue to expand the range of metal
FT	ions the enzyme can use"
FT	Misc-difference 226
FT	/note="Arg of the catalytic domain can be exchanged for
FT	another amino acid residue to expand the range of metal
FT	ions the enzyme can use"
FT	Misc-difference 286
FT	/note="Tyr of the catalytic domain can be exchanged for
FT	another amino acid residue to expand the range of metal
FT	ions the enzyme can use"
FT	Misc-difference 341
FT	/note="Wet of the catalytic domain can be exchanged for
FT	another amino acid residue to expand the range of metal
FT	ions the enzyme can use"
XX	
PN	WO2005056783-A1.
XX	
PD	23-JUN-2005.
XX	
PF	06-DEC-2004; 2004WO-US040844.
XX	
PR	05-DEC-2003; 2003US-0527615P.
XX	
PA	(USSH) US DEPT OF HEALTH.
XX	
P1	Qasba P, Boeggeman E, Ramakrishnan B;
XX	
DR	WPI; 2005-467031/47.
XX	
PT	New catalytic domain from beta(1,4)-galactosyltransferase I that
PT	catalyzes formation of bonds of galactose-beta(1,4)-N-acetylglucosamine
FT	for synthesizing a galactose-beta(1,4)-N-acetylglucosamine group.
XX	
P5	Disclosure; SEQ ID NO 5; 103pp; English.
XX	
XX	The invention describes a new purified and isolated catalytic domain (I)
CC	from a beta (1,4)-galactosyltransferase I that catalyzes formation of a
CC	galactose- beta (1,4)-N-acetylglucosamine bond, glucose- beta (1,4)-N-
CC	acetylglucosamine bond, N-acetylglactosamine- beta (1,4)-N-
CC	acetylglucosamine bond, N-acetylglucosamine- beta (1,4)-N-
CC	acetylglucosamine bond, mannose- beta (1,4)-N-acetylglucosamine bond or
CC	galactose- beta (1,4)-N-acetylglucosamine-6-SO 3 bond in the presence of
CC	magnesium, or N-acetylglactosamine- beta (1,4)-glucose bond in the
CC	presence of alpha-lactalbumin and magnesium. Also described are: a
CC	polypeptide comprising (I); a nucleic acid segment (II) encoding (I); an
CC	expression cassette (III) comprising (II); a cell comprising (II) or
CC	(III); an oligosaccharide comprising a galactose- beta (1,4)-N-
CC	acetylglucosamine group, glucose- beta (1,4)-N-acetylglucosamine group, N-
CC	acetylglactosamine- beta (1,4)-N-acetylglucosamine group, N-
CC	acetylglactosamine- beta (1,4)-glucose group, N-acetylglucosamine- beta
CC	(1,4)-N-acetylglucosamine group, mannose- beta (1,4)-N-acetylglucosamine
CC	group or galactose- beta (1,4)-N-acetylglucosamine-6-SO 3 group
CC	synthesized using (I); increasing the immunogenicity of an antigen using

(I); preparing (M1) a saccharide composition having a defined sequence using (I); a composition comprising (M1); a kit comprising packaging material, and polypeptide comprising (I); and linking a donor into an acceptor that is attached to a blood platelet using (I). (I) Is useful for synthesizing a galactose-beta (1,4)-N-acetylglucosamine group, glucose-beta (1,4)-N-acetylglucosamine group, N-acetylgalactosamine-beta (1,4)-N-acetylglucosamine group, N-acetylgalactosamine-beta (1,4)-N-acetylglucosamine-beta (1,4)-N-acetylglucosamine group, mannose-beta (1,4)-N-acetylglucosamine group and galactose-beta (1,4)-N-acetylglucosamine-6- SO_3 group, which involves incubating a reaction mixture comprising (I), with a donor and an acceptor. (I) Is useful for increasing the immunogenicity of antigen and for preparing a saccharide composition having a defined sequence. (I) Is useful in the preparation of vaccines. This is the amino acid sequence of mouse beta(1,4)-galactosidase I.

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2006, 15:16:49 ; Search time 46 Seconds
(without alignments)
688.365 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSILVLCV.....VDLEKDCRRRLRDPFPCF 383

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5 COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6 COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/ECTUS COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	717	35.3	481	2	US-09-949-016-8746 Sequence 8746, Ap
2	710	35.0	393	2	US-09-118-464-5 Sequence 5, Appli
3	710	35.0	393	2	US-10-132-652A-5 Sequence 5, Appli
4	704	34.7	336	2	US-09-270-767-43058 Sequence 43058, A
5	680.5	33.5	362	1	US-09-055-097-7 Sequence 7, Appli
6	680.5	33.5	362	2	US-09-373-902-7 Sequence 7, Appli
7	673.5	33.2	362	2	US-09-118-464-6 Sequence 6, Appli
8	673.5	33.2	362	2	US-10-132-652A-6 Sequence 6, Appli
9	668.5	32.9	767	1	US-08-446-777-8 Sequence 8, Appli
10	666.5	32.8	767	1	US-08-446-777-6 Sequence 6, Appli
11	666	32.8	398	1	US-08-446-777-2 Sequence 2, Appli
12	657	32.3	372	2	US-09-118-464-2 Sequence 2, Appli
13	657	32.3	372	2	US-09-949-016-8066 Sequence 8066, Ap
14	657	32.3	372	2	US-09-949-016-8067 Sequence 8067, Ap
15	657	32.3	372	2	US-10-132-652A-2 Sequence 2, Appli
16	649	32.0	400	2	US-09-118-464-4 Sequence 4, Appli
17	649	32.0	400	2	US-10-132-652A-4 Sequence 4, Appli
18	646	31.8	342	2	US-09-118-464-3 Sequence 3, Appli
19	646	31.8	344	1	US-10-132-652A-3 Sequence 3, Appli
20	642.5	31.6	344	1	US-09-055-097-3 Sequence 3, Appli
21	642.5	31.6	344	2	US-09-373-902-3 Sequence 3, Appli
22	642.5	31.6	344	2	US-09-949-016-6412 Sequence 6412, Ap
23	642.5	31.6	344	2	US-09-991-181-236 Sequence 236, App
24	642.5	31.6	344	2	US-09-990-444-236 Sequence 236, App
25	642.5	31.6	344	2	US-09-997-333-236 Sequence 236, App
26	642.5	31.6	344	2	US-09-992-598-236 Sequence 236, App
27	642.5	31.6	356	2	US-09-949-016-7638 Sequence 7638, Ap

28	583	28.7	351	2	US-09-949-016-10246 Sequence 10246, A
29	583	28.7	388	2	US-09-949-016-6551 Sequence 6551, Ap
30	577	28.4	382	2	US-09-949-016-6550 Sequence 6550, Ap
31	576	28.4	404	2	US-09-949-016-8230 Sequence 8230, Ap
32	510.5	25.1	164	2	US-09-270-767-58392 Sequence 58392, A
33	314.5	15.5	166	2	US-09-270-767-33562 Sequence 33562, A
34	314.5	15.5	166	2	US-09-270-767-48779 Sequence 48779, A
35	272.5	13.4	327	2	US-09-247-155-94 Sequence 94, Appl
36	272.5	13.4	327	2	US-09-599-3608-12 Sequence 12, Appl
37	272.5	13.4	327	2	US-09-513-989C-12 Sequence 12, Appl
38	272.5	13.4	327	2	US-09-471-276-12 Sequence 12, Appl
39	272.5	13.4	327	2	US-10-012-231A-17 Sequence 17, Appl
40	272.5	13.4	327	2	US-10-015-389A-17 Sequence 17, Appl
41	272.5	13.4	327	2	US-09-903-190-94 Sequence 94, Appl
42	272.5	13.4	327	2	US-10-006-768A-17 Sequence 17, Appl
43	272.5	13.4	327	2	US-10-015-671A-17 Sequence 17, Appl
44	272.5	13.4	327	2	US-10-015-393A-17 Sequence 17, Appl
45	272.5	13.4	327	2	US-10-011-833A-17 Sequence 17, Appl

ALIGNMENTS

```
RESULT 1
US-09-949-016-8746
; Sequence 8746, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1601307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8746
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8746
Query Match          35.3%; Score 717; DB 2; Length 481;
Best Local Similarity 48.9%; Pred. No. 2.2e-72;
Matches 134; Conservative 43; Mismatches 95; Indels 2; Gaps 2;
95 PDNQTQFCNCTPPLVGPRIY-FLDEPDFLTKELYDPTAGHGMPKDCVARRHVAII 153
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
157 PAPQCLPYCPRSPLVGPVSFSPVSLAIYERNRVRVPGRGYRAGEPRSRRTAII 216
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
154 VPIRREHRLIMLNHLSLAKOQDYAIFIVEGNAVQNTNRGGLNMGVGVASRLYPW 213
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
217 VPIRREHRLIMLNHLSLAKOQDYAIFIVEGNAVQNTNRGGLNMGVGVASRLYPW 276
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
214 QCFIHHVDLLPEDDRNLYTC-PIOPRMSVAIDKFNYKLPSYALFGISALTQDLTKI 272
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
277 DCLFLHVDLLPEDDRNLYTC-PIOPRMSVAIDKFNYKLPSYALFGISALTQDLTKI 336
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
273 NGFSNDFWGGGDDDLATRTSMAGLKYSRYPTQIARYKMTIKHSTTEATNPVNRKYTMG 332
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
337 NGFPRYWGCGGDDDLATRTSMAGLKYSRYPTQIARYKMTIKHSTTEATNPVNRKYTMG 396
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
333 QTKRWTRDGLSNLYKYLWNLKPLTYRAVVDL 366
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
397 KTONSWTDGMSLTYQLARELGPLYTNITADI 430
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
```


APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,097
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0490 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 845-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1469908
US-09-055-097-7

Query Match 33.5%; Score 680.5; DB 1; Length 362;

Best Local Similarity 43.9%; Pred. No. 1,9e-68;
Matches 122; Conservative 58; Mismatches 97; Indels 1; Gaps 1;

QY 90 RPILEPDNQTLOFCNQTPPHLVGPRIRVLPDEP-DEKTELEKTYPDTHAGHGMPKDCVARRH 148
DB 80 QPRNLPDSAPSGCLCPDPSPILVGPRIRVFSPGVNLEEVASTNPBVRBGRFAPKCKALQ 139
QY 149 RVAIIIVPRDRBAHLRIMLHNHSLAKQOLDYALFIVEQVANQTFNRGKLMNVGYDVA 208
DB 140 KVAIIIPRNREBEHLKMYLYMHPILOQOLDYGVYVINOQDDEBFNRAKILNVGFTAL 199
QY 209 RLYPQCIFIFHDVLLPEDDRNLYTCPIOPRMSVAIDKFNKLPYSAIFGISALYTDH 268
DB 200 KEYVDVCFVSDVDLIPMDDRNTYKCYOPRHLISVSMKFGRLRYNQGVSALSKEQ 259
QY 269 LKINGFANDPWGCGEDDDLATRTSMAGLKVSRYPTQIARYKMIKHSTEAATPNVKNCR 328
DB 260 FTKINGFPNNYWGCGEDDDLYNRLVFKGMGISRPDAVIGKCRMIHRSDRKNBPNPERF 319
QY 329 KIMGQTKRWRTRDGLSNLKYKLVNLEKPLTYTRAVVDL 366
DB 320 DRIATHTRETMSDGLNSLSYEVLRTDRPPLTYTRITVDI 357

RESULT 6
US-09-373-902-7
Sequence 7, Application US/09373902
Patent No. 6649737
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.

Guegler, Karl J.
Corley, Neil C.
Shah, Purvi
Patterson, Chandra
TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/373,902
FILING DATE: 12-Aug-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/055,097
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0490 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 362 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1469908
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-373-902-7

Query Match 33.5%; Score 680.5; DB 2; Length 362;

Best Local Similarity 43.9%; Pred. No. 1,9e-68;
Matches 122; Conservative 58; Mismatches 97; Indels 1; Gaps 1;

QY 90 RPILEPDNQTLOFCNQTPPHLVGPRIRVLPDEP-DEKTELEKTYPDTHAGHGMPKDCVARRH 148
DB 80 QPRNLPDSAPSGCLCPDPSPILVGPRIRVFSPGVNLEEVASTNPBVRBGRFAPKCKALQ 139
QY 149 RVAIIIVPRDRBAHLRIMLHNHSLAKQOLDYALFIVEQVANQTFNRGKLMNVGYDVA 208
DB 140 KVAIIIPRNREBEHLKMYLYMHPILOQOLDYGVYVINOQDDEBFNRAKILNVGFTAL 199
QY 209 RLYPQCIFIFHDVLLPEDDRNLYTCPIOPRMSVAIDKFNKLPYSAIFGISALYTDH 268
DB 200 KEYVDVCFVSDVDLIPMDDRNTYKCYOPRHLISVSMKFGRLRYNQGVSALSKEQ 259
QY 269 LKINGFANDPWGCGEDDDLATRTSMAGLKVSRYPTQIARYKMIKHSTEAATPNVKNCR 328
DB 260 FTKINGFPNNYWGCGEDDDLYNRLVFKGMGISRPDAVIGKCRMIHRSDRKNBPNPERF 319
QY 329 KIMGQTKRWRTRDGLSNLKYKLVNLEKPLTYTRAVVDL 366
DB 320 DRIATHTRETMSDGLNSLSYEVLRTDRPPLTYTRITVDI 357

RESULT 7
US-09-118-464-6
Sequence 6, Application US/09118464A
Patent No. 6558934

```

; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Bennett, Eric Paul
; TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
; TITLE OF INVENTION: Beta-1,4-Galactosyltransferase, Beta4gal-T2
; FILE REFERENCE: 4305/OE521
; CURRENT APPLICATION NUMBER: US/09/118,464A
; PRIORITY FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
; US-09-118-464-6

Query Match      33.2%; Score 673.5; DB 2; Length 362;
Best Local Similarity 43.5%; Pred. No. 1.2e-67;
Matches 121; Conservative 58; Mismatches 98; Indels 1; Gaps 1;
```

```

; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Bennett, Eric P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; TITLE OF INVENTION: 4-GALACTOSYLTRANSFERASE, B4 GAL-T2
; FILE REFERENCE: 04305/100521-US1
; CURRENT APPLICATION NUMBER: US/10/132,652A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 09/118,464
; PRIORITY FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
; US-10-132-652A-6

Query Match      33.2%; Score 673.5; DB 2; Length 362;
Best Local Similarity 43.5%; Pred. No. 1.2e-67;
Matches 121; Conservative 58; Mismatches 98; Indels 1; Gaps 1;
```

```

; GENERAL INFORMATION:
; APPLICANT: CLAUSER, HENRIK
; APPLICANT: BENNETT, ERIC P.
; TITLE OF INVENTION: UDP-GALACTOSE: B-N ACETYL-GLUCOSAMINE B-1,
; TITLE OF INVENTION: 4-GALACTOSYLTRANSFERASE, B4 GAL-T2
; FILE REFERENCE: 04305/100521-US1
; CURRENT APPLICATION NUMBER: US/10/132,652A
; CURRENT FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 09/118,464
; PRIORITY FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
; US-10-132-652A-6

Query Match      33.2%; Score 673.5; DB 2; Length 362;
Best Local Similarity 43.5%; Pred. No. 1.2e-67;
Matches 121; Conservative 58; Mismatches 98; Indels 1; Gaps 1;
```

```

; GENERAL INFORMATION:
; APPLICANT: Matzele, Eric G.
; APPLICANT: Iwanow, Svetoslav X.
; TITLE OF INVENTION: Proteins having glycosyltransferase
; TITLE OF INVENTION: activity
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: CIBA-GEIGY Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,777
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03194
; FILING DATE: 15 NOV 93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 92810924.8
; FILING DATE: 27 NOV 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19361/A/BE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-777-8

Query Match      32.9%; Score 668.5; DB 1; Length 767;
Best Local Similarity 35.4%; Pred. No. 1.5e-66;
Matches 140; Conservative 74; Mismatches 132; Indels 49; Gaps 6;
```

```

; GENERAL INFORMATION:
; APPLICANT: Matzele, Eric G.
; APPLICANT: Iwanow, Svetoslav X.
; TITLE OF INVENTION: Proteins having glycosyltransferase
; TITLE OF INVENTION: activity
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSER: CIBA-GEIGY Corporation
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: NY
; ZIP: 10591-9005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,777
; FILING DATE: May 26, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP93/03194
; FILING DATE: 15 NOV 93
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EPO 92810924.8
; FILING DATE: 27 NOV 92
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 4-19361/A/BE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 277-3318
; TELEFAX: (908) 277-4306
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 767 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-777-8

Query Match      32.9%; Score 668.5; DB 1; Length 767;
Best Local Similarity 35.4%; Pred. No. 1.5e-66;
Matches 140; Conservative 74; Mismatches 132; Indels 49; Gaps 6;
```


MOLECULE TYPE: protein
US-08-446-777-2

Query Match 32.8%; Score 666; DB 1; Length 398;
Best Local Similarity 35.7%; Pred. No. 1e-66;
Matches 137; Conservative 73; Mismatches 128; Indels 46; Gaps 5;

14 LVLGCVLLVHMIT-----KIPSYENLTIGSTLADVDAMEAVLGNFA-----60
26 LVAVCALHLGVTLYVLAGRDLRLPOL-----VGSTPLGGSSNAAAGSSGGLRIG 80
QY 61 -----STSDLLDTWNSIFSPISEVNOTSFMEDIRPILFPDNOQLQFC 103
DB 81 GARPPEPLGASQPRPGDSSPVVDGSPASMLTSY-----PVPHTLALSIPAC 130
QY 104 NQPPHNLGIRFLEDP-DFKLEKTYRPTHAGGCMPCDVARHRAIIVPYRDEAH 162
DB 131 PEESPILVGMLEFNNPVLELVARQNPVKMGGRYAPRDCVSPHKVAIIFRRQEH 190
QY 163 LRIMLNHLSLAKQOLDYAIFIVEGVANOTFNRGKLMNGYDVASRLPYMOGCFIHVDV 222
DB 191 LKWTLYLHVLROQDLDIYIYNQAGDTIFNRKALINVGPEALKDYITCFVFSVDV 250
QY 223 LLEPDDRNLYTCPIOPRHMSVAIDKENYKLPYSAIFGIGISALTGDHLKINGFSDNFGW 282
DB 251 LIPMNDHNAVRCFSQPRHISVAMDKFGFSLPYQVFGVSALSKQOFLTINGFPNNYMGW 310
QY 283 GGGDDDLARTSMAGLKVSRYPQIARYKMKIKSTEAITNVNCKRYKIMQTRKRTROG 342
DB 311 GGGDDDI FNNLVFRGMSISRPNAVGVGCRMIIRSRDKNEPNPFRDLIAKTEMLSDG 370
QY 343 LSNLYKLVNLEKPLTYTRAVVDL 366
DB 371 LNSLTYQVDLVQRYRPLTYQTIVDI 394

RESULT 12
US-09-118-464-2
Sequence 2, Application US/09118464A
Patent No. 6558934
GENERAL INFORMATION:
APPLICANT: Clausen, Henrik
APPLICANT: Bennett, Eric Paul
TITLE OF INVENTION: UDP-Galactose: Beta-N-Acetyl-Glucosamine
TITLE OF INVENTION: Beta-1,4-Galactosyltransferase, Beta4gal-12
FILE REFERENCE: 4305/08521
CURRENT APPLICATION NUMBER: US/09/118,464A
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 372
TYPE: PRT
ORGANISM: Homo sapiens
US-09-118-464-2

Query Match 32.3%; Score 657; DB 2; Length 372;
Best Local Similarity 37.2%; Pred. No. 9.6e-66;
Matches 135; Conservative 71; Mismatches 135; Indels 22; Gaps 6;

12 KSLVLVCAVLVLAHMIYKIPSLYENLTIGSS-TLADVDAMEAVLGNFASDSDLDTW 70
DB 14 KAVLLLCLLHFLVAVILY-FDYVAQHLAFPSRSGARPAHALHPAASSSSSSCSRPNA 72
QY 71 NSTFSPISEVNOTSFMEDIRPILFPDNOQLQFCNQTPPHLVGPIRVFLDEP-DFKTIEXI 129
DB 73 TASSSGLPYVPSA-----LPGPTAPTLPPCPDSPPGVLGRLLIFTSFMPLEVRQRE 124
QY 130 YPPTHAGGCMPCDVARHRAIIVPYRDEAHRLIMLNHLSLAKQOLDYAIFIVEGV 189
DB 125 NPGVLMGGRYTPPDCPTPAQTVAVIIPFRHREHLRYMLHPLIRORLRKYGVIYNQH 184
QY 190 ANQTFNRGKLMNNGY-----DVASRLPYMOGCFIHVDVLLPEDDRNLYTCPIOPRHMSV 243

185 GEDTFNRKALINVGFILEALKEDAA-----YDCFISVDVLMVMDRNLYRCGDQPRHFAI 239
QY 244 AIDKFNYKLPYSAIFGIGISALTGDHLKINGFSDNFGMGEGDDDLATRTSMAGLKVSR 303
DB 240 AIDKFGFRLPYAGYFGVSGLSKAQFLRINGFPNFWMGSGDDDI FNNISLTGMKISR 299
QY 304 PQIARYKMKIKSTEAITNVNCKRYKIMQTRKRTROGSLNLYKLVNLEKPLTYTRAV 363
DB 300 DIRIGRYMKIKDRDKHNEPNPQRTKIONTKLTMRKDGIGISVRYQVLEVSQPLFTNIT 359
QY 364 VDL 366
DB 360 VDI 362

RESULT 13
US-09-949-016-8066
Sequence 8066, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8066
LENGTH: 372
TYPE: PRT
ORGANISM: Human
US-09-949-016-8066

Query Match 32.3%; Score 657; DB 2; Length 372;
Best Local Similarity 37.2%; Pred. No. 9.6e-66;
Matches 135; Conservative 71; Mismatches 135; Indels 22; Gaps 6;

12 KSLVLVCAVLVLAHMIYKIPSLYENLTIGSS-TLADVDAMEAVLGNFASDSDLDTW 70
DB 14 KAVLLLCLLHFLVAVILY-FDYVAQHLAFPSRSGARPAHALHPAASSSSSSCSRPNA 72
QY 71 NSTFSPISEVNOTSFMEDIRPILFPDNOQLQFCNQTPPHLVGPIRVFLDEP-DFKTIEXI 129
DB 73 TASSSGLPYVPSA-----LPGPTAPTLPPCPDSPPGVLGRLLIFTSFMPLEVRQRE 124
QY 130 YPPTHAGGCMPCDVARHRAIIVPYRDEAHRLIMLNHLSLAKQOLDYAIFIVEGV 189
DB 125 NPGVLMGGRYTPPDCPTPAQTVAVIIPFRHREHLRYMLHPLIRORLRKYGVIYNQH 184
QY 190 ANQTFNRGKLMNNGY-----DVASRLPYMOGCFIHVDVLLPEDDRNLYTCPIOPRHMSV 243
DB 185 GEDTFNRKALINVGFILEALKEDAA-----YDCFISVDVLMVMDRNLYRCGDQPRHFAI 239
QY 244 AIDKFNYKLPYSAIFGIGISALTGDHLKINGFSDNFGMGEGDDDLATRTSMAGLKVSR 303
DB 240 AIDKFGFRLPYAGYFGVSGLSKAQFLRINGFPNFWMGSGDDDI FNNISLTGMKISR 299
QY 304 PQIARYKMKIKSTEAITNVNCKRYKIMQTRKRTROGSLNLYKLVNLEKPLTYTRAV 363
DB 300 DIRIGRYMKIKDRDKHNEPNPQRTKIONTKLTMRKDGIGISVRYQVLEVSQPLFTNIT 359
QY 364 VDL 366
DB 360 VDI 362

THIS PAGE BLANK (USPTO)

Qy	422	CCTAAGAGATTTGGTTGCAAGCATCGTGTCTTATTTATGGCCCTTATAGAGATCGTGA	480
Db	461	CCTCAGGAATGTAAAGCTTTTACAGAGGGTCCCATCTCGTTCCCAACGGAAACAGAGAG	520
Qy	481	GCACATTTGAGAAATTAAGTCTCCACAATTTGCACTCGTTGCTCGGCAACAACAATTGGAC	540
Db	521	AAACACTGATGTACTCTGCTGGAACATCTGCATCCCTTCTGCAAGGACACAGCTGGAT	580
Qy	541	TATGCAATTTTCATTGTGGAGCAGCATGGCGAATCAAGCGTTTAATCGCGGAAACTATG	600
Db	581	TATGGATCTACGTATCATCCACAGGCTGAAAGTAAAGATTATTCAGAGCCAAACTCTTG	640

QY	601	AACGTGGATACGACGTATGACATCAACGCTTACACCATGGGAGTGCCTCATTTTCAATGAT	660
Db	641	AATGTGGCTATCTTAGAAGCCCTCAGAAGAAATTTGGACTGCTTTATATTTCCAGAT	700
QY	661	GTGCATTTACTGCCCGAAGATGACCGTAACTGTACAGTGTCCAATTTCAACACGTCAT	720
Db	701	GTGACACCTGATACCCGAGAAATGACTTTTAACCTTTTAAAGTGTGAGAGCATCCCAACAT	760
QY	721	ATGAGTGTAGCGATGCATTAATTTCAATATATTAATCTTCATATTTGGGCGATCTTGGGCGGA	780
Db	761	CTGGTGTGTGGCAGGAACAGCATCTGGTATCAAGTTACGTTACAGTGAATATTTGGGGGT	820
QY	781	ATCAGTGCATTACAAAAGATCACTCGTAAGAAAATCAATGSAATTTTGCATGATTTTGG	840
Db	821	GTTACTGCGCTTAGACGAGACAGACTTTTTCAAAGTGATGATTCCTTAACAACTACTCG	880
QY	841	GGTTGGGGCGGAGGAGCAGCATTTTGGCGACGAAATCATGAGTGGCTGAGCTGAAGTT	900
Db	881	GGATGGGAGGCGAGACGATGACTCAGACTCAGGGTTGAGCTTCAAAGATTAATTT	940
QY	901	TCAAGATATCCGACACAAATTTGCACGATATTAATATGATTGAACACTGCAGCGAAGCAGC	960
Db	941	TCCGGGCCCCCTGCGTAGTGGGTAAATATATCAATGTGCTTCCACATPAGAGACAAAGGC	1000
QY	961	AATCCAGTTATTAATGCGCGCTACAAATTAATGGCGCCAAAGAGCGCGCATGAGCAAGT	1020
Db	1001	AATGAGTGAACGCAAGACGATGAAGCTCTTACACCAAGTGTACAGAGTCTGGAGAAC	1060
QY	1021	GACGGCCTAGCAATCTGAAGTATAGCTCGTAAATCTGGAATGGAAGCCTCTTACACT	1080
Db	1061	GATGGGTTGAGATGTTGTTCTTATTAATTAAGATCTGTGAAACAACTCTTATATATTC	1120
QY	1081	CGAGCGCTGTCGATTT	1097
Db	1121	AACATCACAGTGGATTT	1137

```

RESULT 2
US-10-955-054A-120
; Sequence 120, Application US/10955054A
; Publication No. US20050266420A1
; GENERAL INFORMATION:
; APPLICANT: PUSZTAI, LAJOS
; APPLICANT: SYMMANS, W. FRASER
; APPLICANT: HESS, KENNETH R.
; APPLICANT: AYERS, MARK
; APPLICANT: STEC, JAMES
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY
; FILE REFERENCE: UTXC:880US
; CURRENT APPLICATION NUMBER: US/10/955,054A
; CURRENT FILING DATE: 2004-09-30
; NUMBER OF SEQ ID NOS: 195
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 120
; LENGTH: 4646
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-955-054A-120

Query Match      13.0%; Score 150; DB 6; Length 4646;
Best Local Similarity 56.8%; Pred. No. 4,1e-40;
Matches 276; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

405 CGGTGACATGGAATGGCTTAAGATTGGTTGCAAGGCACTGCGTGTGCTATTATTGTC 464
559 CGAAGTCTACTGGAAGCCTTCTGATTGATGCGCTCGTGGAAGGTGGGATCTTATCC 618
465 CTATAGAGATCGTGAAGCATTTTGAATATATGCTCCACATTTTGCACTCGTTCCTGC 524
619 CTTCCGAACCGCCACAGACACCTCCAGTCTGTTTCAAGACACCTGCTCCATCTCCA 678
525 CAACCAACATTGGACTATGCAATTTTCAATTGTGAGCAAGTGGCGAATCAGACGTTTA 584

```

Db	679	GGCCAGCGGCTGGAGTTTGATTTTNGTGTGAACAAGTTGTAACCCAAACCCCTTTAA	738
QY	585	TCGGGGGAAACTAATGAACGTTGGATACGAGCTGATCAAGCCTTCATACCATGGCAGTG	644
Db	739	TCGAGCCATGCTTTTTCACGTTGGCTTTCAAGAGGCATGAAGACCTTGGATTGGGACTG	798
QY	645	CTTATCATCTTCATGATGTCGATTTACTGCGCCGAGATGACCGTAACTGTACAGTGTCC	704
Db	799	TTTGATTTTTCATATGATGTAGATCAATACCGGAAAGTATGCGAATTTATGTGATGTGG	858
QY	705	AATTCAACCAAGTCATATGAGTGTAGCGATCGATTAATTCAATTATAAACCCTTCATATTC	764
Db	859	ACAGATGCCAGGAGCATTTTGGCAACCAATGATGATAGATATGTATCTGCTTCCCTTATAC	918
QY	765	GGCGATCTTCGGCGGGAATCAGTGCATCAACAAAAGATCACTGAAGAAAATCAATGATGATT	824
Db	919	CGAGTTCCTTTTGGCGGAGTGAAGTGGCTTAAACAGTGGAAACAATTTTGGAAAAATCAATGGCTT	978
QY	825	TTTGAATGATTTTGGGGTTGGGGCGGAGAGGACGAGATTTTGGCGACGAGAACATCGAT	884
Db	979	TCCCAATGCTTTTCGGGGTGGGGGTGGGAAGATGACGACCTCTTGGAACAAGATCACGAA	1038
QY	885	GGCTGG	890
Db	1039	TGCAAG	1044

```

US-11-136-527-1055
/ RESULT 3
/ Sequence 1055, Application US/11136527
/ Publication No. US20050287570A1
/ GENERAL INFORMATION:
/ APPLICANT: Wyeth
/ APPLICANT: Mounts, William M
/ TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
/ FILE REFERENCE: 031896-041000 (AM101068)
/ CURRENT APPLICATION NUMBER: US/11/136,527
/ CURRENT FILING DATE: 2005-05-25
/ PRIOR APPLICATION NUMBER: US 60/574,294
/ PRIOR FILING DATE: 2005-05-26
/ NUMBER OF SEQ ID NOS: 362830
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1055
/ LENGTH: 951
/ TYPE: DNA
/ ORGANISM: Rattus norvegicus
US-11-136-527-1055

Query Match      12.0%; Score 138.8; DB 7; Length 951;
Best Local Similarity 48.5%; Prid. No. 9.8e-37;
Matches 390; Conservative 7; Mismatches 404; Indels 3; Gaps 14;

QY      294 GACCTTGAATTCTGTATCGACAGACCCGCCACCTCGTGGAGACCATCCGNGATTCCT 353
      |||||
DB      132 GACTGTGCTGCTTGCTCCCTGAGAGACTCCCGCTGCTCGTTGGYCCCATGCGATTGACTT 191
      |||||

QY      354 CGATGAGCC---CGACTTCAAAACTCTCGAGAAATCTATCCGACACGACGCCGCTGG 410
      |||||
DB      132 TAAATTTCTCTGTGATCTGAGAGCTTTTGGCAAGAAAGAACCAKGAATAAAGATGGCGG 251
      |||||

QY      411 ACATGGAATGCTTAAAGATGTGTGTCGAAGCATGCTGTGCTATATTATGTCCTTAAAG 470
      |||||
DB      252 SCGTACTTCCCAAGGCMCTGATCTCCCTCACAAAGGTGGCCATCATTAATCCCATTCG 311
      |||||

QY      471 AGATCGTGAAGACATTTGAGATAATGCTCCACAATTTGGACGTGTTGCTCGCCAAACA 530
      |||||
DB      312 TAACCGGAGAGACCTCTCAAAATACTGCTGTATTATTGCAATCCAGTCTTCAGCGCA 371
      |||||

QY      531 ACAATTGACTATGCAATTTTTCATTGTGAGAGCACTGGCGAATCGATCAGAGCTTTAATCGCG 590
      |||||
DB      372 GCAACTGCACTATGGCATCTAGTCATCAATCAGGCTGGAAGACACCATGTTTAATCGAGC 431
      |||||

QY      591 GAACTAATGAACGTTGGATACAGCTAGCATCAGCCTCTTAACCATGGCAGTCTTCAT 650
      |||||

```

```

Db 432 TAAAGTGCTCAACGTTGGCTTTCAGAGAGCCCTTGAAGACTATGACTAAGCTGGTTGT 491
Qy 651 CTTTCATGATGTCGATTTTACTGCCGAGATGACCGTAACTGTACAGCTGTCCAAATTC 710
Db 492 GTTCAGTATGTCGATTTTACTGCCGAGATGACCGTAACTGTACAGCTGTCCAAATTC 710
Qy 711 ACCAGTCAATGAGTATGTCGATTTTACTGCCGAGATGACCGTAACTGTACAGCTGTCC 770
Db 552 GCGACGCGATATTTCTGTGCAATGACGATTTGGGTTTACCTGCTTACGTTCAATTA 611
Qy 771 CTTCCGCGGAATCAGTGCATTAACAAAGATCAGCTGAAGAAATCAATGATTTTCGA 830
Db 612 TTTTGAAGGTGTCTCCGTTCTGATTAACAGAGTTTCTTACCATTAATGATTTTCGA 671
Qy 831 TGAATTTTGGGCTGGGCGGAGAGACGACATTTTGGCAGAGAACATTCATGAGCTGG 890
Db 672 TAATTACTGGGCTGGGCGGAGAGATGATGATTTTAAACAGATTTGTTCAATTAAG 731
Qy 891 ACTGAAGTTTCAAGTATTCGACACAAATTCGATTAATTAATGATTAACGATCGAC 950
Db 732 CATGCTATATACCGCCCAATGCTGTGTAGGCGRGTGTGCAATGATCGGACACTCAAG 791
Qy 951 GGAAGCGACGATCCAGTTAATTAATGCGCTAACAATAATGAGCGCAACGAGCGCG 1010
Db 792 AGACAAGTTTATGAGCCCAACCTTCAGAGTTTGAACCGGATCGCACATTAAGAAAC 851
Qy 1011 ATGACACGTCGACGCGCTTAAGCAATCTGAATTAAGCTGTAATCTGAAATTTGAAGC 1070
Db 852 GATGCGCTTGAATGTTTGAATCTACTACCTACGAGGTGTGACATACAGATGATACC 911
Qy 1071 TCTTACACTCGAGCGCTGTGCA 1094
Db 912 GTTATATACCAAAATTCACAGTGA 935

```

RESULT 4

```

US-11-136-527-5151
; Sequence 5151, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5151
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-5151

```

```

Query Match 12.0%; Score 138.8; DB 7; Length 951;
Best Local Similarity 48.5%; Pred. No. 9, 8e-37;
Matches 390; Conservative 7; Mismatches 404; Indels 3; Gaps 1;

```

```

Qy 294 GACTCTTCAATTTCTGTAATCAAGACACTTCCCACTGTGCGAGCCATCCGTATTTCT 353
Db 132 GACTCTGCTGCTTGTGCTGAGAGATCCCGCTGCTGTGCTGCTGCTGCTGCTGCTGCT 191
Qy 354 CGATGAGCC---CGACTTCAAACTCTGAGAAATCTTCCGAGACACGACGCGGTG 410
Db 192 TATATATCTGTGATCTGAGCTTTTGGCAAGAAAGAACCAKATTAATGAGGCGG 251
Qy 411 ACATGGAATGCTAAGATTTGTGCAAGGACATCGTTGCTATTAATTTGCTTATAG 470
Db 252 SGTTACTTCCCAAGMCTGTATCTCCCTCAAGAGTGGCCATCATTAATCCATTCG 311

```

```

Qy 471 AGATGTAAGACACATTTGAGAAATATGCTCCACATTTGCACTGCTGCGCAACA 530
Db 312 TAACCGGAGAGACACCTCAAAATCTGGCTGATTAATTTGATTCAGTCTTCAAGCGCA 371
Qy 531 ACAATGGAATTAAGATTTTCAATTTGTCGAGCAATGCGCAATCAGAGTTTAAATCGCG 590
Db 372 GCAATCTGATGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
Qy 591 GAAATTAATGAACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 650
Db 432 TAAAGTGTCAACGTTGGCTTTCAAGAGCGCTTGAAGACTATGATCAATGCTTTGT 491
Qy 491 CTTTCATGATGTCGATTTTACTGCCGAGATGACCGTAACTGTACAGCTGTCCAAATTC 710
Db 651 GTTCAGTATGTCGATTTTACTGCCGAGATGACCGTAACTGTACAGCTGTCCAAATTC 671
Qy 711 ACCAGTCAATGAGTATGTCGATTTTACTGCCGAGATGACCGTAACTGTACAGCTGTCC 770
Db 552 GCGACGCGATATTTCTGTGCAATGACGATTTGGGTTTACCTGCTTACGTTCAATTA 611
Qy 771 CTTCCGCGGAATCAGTGCATTAACAAAGATCAGCTGAAGAAATCAATGATTTTCGA 830
Db 612 TTTTGAAGGTGTCTCCGTTCTGATTAACAGAGTTTCTTACCATTAATGATTTTCGA 671
Qy 831 TGAATTTTGGGCTGGGCGGAGAGACGACATTTTGGCAGAGAACATTCATGAGCTGG 890
Db 672 TAATTACTGGGCTGGGCGGAGAGATGATGATTTTAAACAGATTTAGTTCAATTAAG 731
Qy 891 ACTGAAGTTTCAAGTATTCGACACAAATTCGATTAATTAATGATTAACGATCGAC 950
Db 732 CATGCTATATACCGCCCAATGCTGTGTAGGCGRGTGTGCAATGATCGGACACTCAAG 791
Qy 951 GGAAGCGACGATCCAGTTAATTAATGCGCTAACAATAATGAGCGCAACGAGCGCG 1010
Db 792 AGACAAGTTTATGAGCCCAACCTTCAGAGTTTGAACCGGATCGCACATTAAGAAAC 851
Qy 1011 ATGACACGTCGACGCGCTTAAGCAATCTGAATTAAGCTGTAATCTGAAATTTGAAGC 1070
Db 852 GATGCGCTTGAATGTTTGAATCTACTACCTACGAGGTGTGACATACAGATGATACC 911
Qy 1071 TCTTACACTCGAGCGCTGTGCA 1094
Db 912 GTTATATACCAAAATTCACAGTGA 935

```

RESULT 5

```

US-10-995-561-13225/c
; Sequence 13225, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: Carigill, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; FILE REFERENCE: CLO01559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13225
; LENGTH: 16822
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(16822)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-2)
US-10-995-561-13225

```

```

Query Match 4.6%; Score 53.2; DB 6; Length 16822;
Best Local Similarity 63.1%; Pred. No. 1, 1e-06;
Matches 82; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```

Qy	750	TAAACTTCATATTTTCGGCGATCTTTCGGCGAATCAGTGCATCTACAAAGATCACCTGAA	809
Db	13331	TACGCTCCCTACCCCCCACTGACGAGGAGCTCAGCATTACTCTGACACAGTACT	13272
Qy	810	GAAATCAATGATTTTTCGAATGATTTTTCGGGCTTGGGCGAGAGACGACGATTTGGC	869
Db	13271	GAAAGATGAAGGCTCCCAATGATATTCGGGCGTGGGAGTGAAGATGACGACATTGC	13212
Qy	870	GACGAGAACCA	879
Db	13211	TACGAGTCA	13202

RESULT 6
US-10-995-561-13401

```

? Sequence 13401, Application US/10995561
? Publication No. US20050272054A1
?
? GENERAL INFORMATION:
? APPLICANT: CARGILL, Michele et al.
? TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
? TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
? TITLE OF INVENTION: DETECTION AND USES THEREOF
? FILE REFERENCE: CLO01559
? CURRENT APPLICATION NUMBER: US/10/995,561
? CURRENT FILING DATE: 2004-11-24
? NUMBER OF SEQ ID NOS: 85702
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 13401
? LENGTH: 18705
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)...(18705)
? OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1
? US-10-995-561-13401

```

Query Match 4.6%; Score 53.2; DB 6; Length 18705;

Best Local Similarity 63.1%; Pred. No. 1.2e-06;
Matches 82; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

750 TAACTTCATATTCGGCGATCTTCGGCGGAA TCAGTCACCTAACAAAAGATCCTGAA 809

Db 10281 TAGCCTCCCGTACCCCGAGTACTTTGGAGGAGTCTCAGCATTTACTTCTTGACCACTTCTT 10340

810 GAAATCAATGATTTTCGAATGATTTTGGGTTGGGCGAGAGAGACGACGATTTGGC 869

Db 10341 GAAGATGAATGGCTTCCCAATGAATACTGGGGCTGGGGTGGTGAGGATGACGACATTGC 10400

Qy 870 GACGAGACA 879

RESULT 7

US-10-750-185-30398
Sequence 30398, Application US/10750185
Publication No. US2005026603A1
GENERAL INFORMATION:
APPLICANT: MMT GENOMICS, INC.
APPLICANT: DENISE, Sue K.
APPLICANT: KERR, Richard
APPLICANT: ROSENFELD, David
APPLICANT: HOLM, Tom
APPLICANT: BATES, Stephen
APPLICANT: FANTIN, Dennis
TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
FILE REFERENCE: MM10100-2
CURRENT APPLICATION NUMBER: US/10/750,185
CURRENT FILING DATE: 2003-12-31
PRIOR APPLICATION NUMBER: US 60/437,482
PRIOR FILING DATE: 2002-12-31
NUMBER OF SEQ ID NOS: 64922

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30398
;
; LENGTH: 2507
;
; TYPE: DNA
; ORGANISM: Bovine
;
US-10-750-185-30398

```

Query Match	4.4%	Score 50.8;	DB 6;	Length 2507;
Best Local Similarity	62.7%	Pred. No. 2.2e-06;		
Matches 79;	Conservative	0;	Mismatches 47;	Indels 0;
				Gaps 0;

754 CTTCATATTGGCGATCTTCGGCGGAATCAGTCACCTAACAAAGATCACCCTGAAGAAA 813

Db 1007 CTCCTGATCCCACTTCTGCGGGCTCTCAGCGCTACCTCTGACCACTGTAAG 1066

QY 814 ATCAATGATTTTCGATGATTTTGGGGTTGGGGCGAGAGACGATTTGGCGACG 873

Dd 1067 ATGATGCGCTTCCCAATGAATACTGGGGCTGGGGTGGTGAGGATGACGCAATTGCTACC 1126

QY 874 AGACCA 879

RESULT 8

```

US-10-750-623-30398
; Sequence 30398, Application US/10750623
; Publication NO. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 30398
; LENGTH: 2507
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-623-30398

```

Query Match	Score	DB	Length
4.4%	50.8	6	2507

Best Local Similarity 62.7%; Pred. NO. 2.4e-06;
Matches 79; Conservative 0; Mismatches 47; Indels 0; Gaps 0.

754 CTTCCATATTCGGCGATCTTCGGGGGAATCAGTCACCTAACAAAGATCATCTGAGAAA 813

Db 1007 CTCCCGTATCCCGAGTACTTGGGGGGTCTCAGCGCTCACTCTGACCAGTACTGAG 1066

814 ATCATATGATTTCGAATGATTTTGGGGTGGGGCGGAGAGGACGACCATTTTGGCCGACG 873

Db 1067 ATGATGCGCTTCCCAATGATATCTGGGCGTGGGGTGGTGGAGATGACGACATTGCTACC 1128

QY 874 AGACA 879
|||
1107 10000 1100

RESULT 9

```

US-10-995-561-22169/c
; Sequence 22169, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.

```

```

METHODS OF
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE,
TITLE OF INVENTION: DETECTION AND USES THEREOF
FILE REFERENCE: CL001559
CURRENT APPLICATION NUMBER: US/10/995,561
CURRENT FILING DATE: 2004-11-24
NUMBER OF SEQ. ID NOS: 85702
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 22169
LENGTH: 201
TYPE: DNA
ORGANISM: Homo sapiens
US-10-995-561-22169

```

	Query Match	Similarity	4.4%;	Score 50.2;	DB 61;	Length 201;
	Best Local	Similarity	57.5%;	Pred No. 6.8e-07;		
	Matches	86;	Conservative	1;	Mismatches 64;	Indels 0; Gaps 0
QY	550	TTGATTGTGGAGCACTGCGCATGACAGCGTAAATCGCGGAAAATAATGAAGCTTGGA	609			
Db	154	TTCATTATCTTCAGCGCTGGAAATGGAACATTTAACAAGGAAAACTGTTGAATGTTGGG	95	:		
QY	610	TACGACGTGACATCACGCCTTAACCCANTGCAAGTCTTCATCTTCAATGATGTCGATTYA	669			
Db	94	GTCGCGAAGGCCCTGTGCGTATGAAGAAGTGGGACTCCTGTTCTTGACGATGTGGAACCTC	35			
QY	670	CTGCGCGAAGATGACCGTAACCTGTACACGNT	702			
Db	34	TTGCGAAGAAATGACCACCATCTGTATGTGTGT	2			

```

RESULT 10
US-10-995-561-68468
; Sequence 68468, Application US/10995561
; Publication No. US20050272054A1
;
GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 68468
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-995-561-68468

```

	Query Match	4.4%;	Score 50.2;	DB 6;	Length 201;
	Best Local Similarity	57.5%;	Pred. No. 6.8e-07;		
	Matches	88;	Conservative	1;	Mismatches 64; Indels 0; Gaps 0
QY	550	TTGATTGTGGAGCAAGTGGCGAATCAGACGTTTAATCCGGGAAACTATGACGTTGGA	609		
Db	48	TCATTATCTTCAGGCTGGAATGGAACATTTACAGGGCAAACTGTGAAGTTGGG	107		
QY	610	TACGACGTGACATCAGCGCTTAACCATGGCAGTGCCTCAATCTTCATGATGTCGATTTA	668		
Db	108	GTGCGAGAGGCCCTCGCGTATGATGAGAGTGGGACTCGTCTGTTCTTGACGATGTGACCTC	167		
QY	670	CTGCGGAGATGATGACCGTAACTGTACACGTGT	702		
Db	168	TTGCCAGAAAATGACCAACATCTGATGTGTGT	200		

```

RESULT 11
US-10-750-185-43601
; Sequence 43601, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:

```

```

? APPLICANT: MMI GENOMICS, INC.
? APPLICANT: DENISE, Sue K.
? APPLICANT: KERR, Richard
? APPLICANT: ROSENFELD, David
? APPLICANT: HOLM, Tom
? APPLICANT: BATES, Stephen
? APPLICANT: FATTIN, Dennis
? TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
? FILE REFERENCE: MM1100-2
? CURRENT APPLICATION NUMBER: US/10/750,185
? CURRENT FILING DATE: 2003-12-31
? PRIOR APPLICATION NUMBER: US 60/437,482
? PRIOR FILING DATE: 2002-12-31
? NUMBER OF SEQ ID NOS: 64922
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 43601
? LENGTH: 1728
? TYPE: DNA
? ORGANISM: Bovine 19866880501107
US-10-750-185-43601

```

	Query Match	3.6%	Score 41.8;	DB 6;	Length 1728;
	Best Local Similarity	53.3%	Pred. No. 0.002;		
	Matches	86;	Conservative	0;	Mismatches 77; Indels 0; Gaps 0;
QY	699	GATGCAATTCACACGCTCATATGAGTGTGAGCATGATTAATTCATTATTAACCTCC	758		
Db	1165	GAGAACAAAGAGACTTCGTGCTAACCTTCCTGCTCTCTCCCTTCAGGTACG	1224		
QY	759	ATATTCGGCCATCTTCGGCGGAATCAGTCGACTTAACAAAAGATCACTGAAAGAAATCA	818		
Db	1225	TTAAAGTGAATATTTTGGGGGTGTACTTCGCCCTTAAGTCGAGAGCAATTTTCAAGTGAA	1284		
QY	819	TGATTTTCGATGATTTTGGGGGTGGGGCGGAGAGGACGACGA	863		
Db	1285	TGATTTCTTAACACTACTGCGGATGGGAGCGGAAAGACGATGA	1329		

```

RESULT 12
US-10-750-623--43601
: Sequence 43601, Application US/10750623
: Publication No. US20050287531A1
: GENERAL INFORMATION:
: APPLICANT: MMI GENOMICS, INC.
: APPLICANT: DENISE, Sue K.
: APPLICANT: KERR, Richard
: APPLICANT: ROSENFELD, David
: APPLICANT: HOLM, Tom
: APPLICANT: BATES, Stephen
: APPLICANT: FANTIN, Dennis
: TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
: FILE REFERENCE: MM1100-1
: CURRENT APPLICATION NUMBER: US/10/750,623
: CURRENT FILING DATE: 2003-12-31
: PRIOR APPLICATION NUMBER: US 60/437,482
: PRIOR FILING DATE: 2002-12-31
: NUMBER OF SEQ ID NOS: 64922
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 43601
: LENGTH: 1728
: TYPE: DNA
: ORGANISM: Bovine 19866880501107
US-10-750-623--43601

```

[illegible]

Db 1225 TTACAGTGGATATTTGGGGGTGTACTGCGCTAAGTCGAGACAAATTTTCAAGTGAA 1284
Qy 819 TGGATTTTGAATGATTTTGGGGGTGGGGCGGAGAGACGACGA 863
Db 1285 TGGATTCTTAACAACACTACTGGGATGGGAGGCGAAGACGATGA 1329

RESULT 13

US-10-750-185-34973
; Sequence 34973, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34973
; LENGTH: 3458
; TYPE: DNA
; ORGANISM: Bovine 19866881077545
US-10-750-185-34973

Query Match 3.0%; Score 34.6; DB 6; Length 3458;
Best Local Similarity 67.1%; Pred. No. 0.91;
Matches 49; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 630 CTACCATGCGAGCGTTCATCTTCATGATGCGATTCGCGCGAAGATGACCGTAA 689
Db 3359 CTCACCTTAACAACGCTTCATCTTCAGTACGTGACCTGCTCCCATGATGACCGCA 3418
Qy 690 CCTGTACACGTGT 702
Db 3419 CCTGTACCGCTGT 3431

RESULT 14

US-10-750-623-34973
; Sequence 34973, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34973
; LENGTH: 3458
; TYPE: DNA
; ORGANISM: Bovine 19866881077545
US-10-750-623-34973

Query Match 3.0%; Score 34.6; DB 6; Length 3458;
Best Local Similarity 67.1%; Pred. No. 0.91;
Matches 49; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

Qy 630 CTACCATGCGAGCGTTCATCTTCATGATGCGATTCGCGCGAAGATGACCGTAA 689
Db 3359 CTCACCTTAACAACGCTTCATCTTCAGTACGTGACCTGCTCCCATGATGACCGCA 3418
Qy 690 CCTGTACACGTGT 702
Db 3419 CCTGTACCGCTGT 3431

RESULT 15

US-10-467-657-4833
; Sequence 4833, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqMan99, version 1.04
; SEQ ID NO 4833
; LENGTH: 879
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4833

Query Match 2.9%; Score 33.6; DB 6; Length 879;
Best Local Similarity 57.7%; Pred. No. 0.83;
Matches 60; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 787 GCACCTAACAAGATCACTGAAAGAAATCAATGATTTTCGAATGATTTTGGGGTGG 846
Db 406 GCACCTTTGAAGGAAGCGCCGGAATAATTCATTGTGTTTGGACGATTTGTCATTTGAA 465
Qy 847 GCGCGAGAGACGACGATTTGGGACGAGACATGATGGCTGG 890
Db 466 AGCGGCGACGAAACCTATTAAGGCATTGAAAACGGCATTTGGATGG 509

Search completed: January 9, 2006, 01:00:42
Job time : 282 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 8, 2006, 22:31:00 ; Search time 746 Seconds
(without alignments)
3421.685 Million cell updates/sec

Title: US-10-661-430-1
Perfect score: 2031
Sequence: 1 MAFRLAVARLKSLLVLCANV.....VDLEKDCRRRLRDRPFCF 383

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues
Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV-xlh
-MODEL=frame+ p2n.model -DEV-xlh
-DB=N.Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10661430.GCGN.1.1.727 @runat_06012006.123134.6226 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N.Geneseq.21.*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2031	100.0	1152	12	ADM32969 Nucleotid
2	718	35.4	2002	3	AACT77836 Human can
3	717	35.3	1938	2	AAZ42120 Human and
4	717	35.3	1949	13	ACN37523 Tumour-as

5	709	34.9	1839	4	ABL14209	Ab114209 Drosophila
6	684	33.7	1853	2	AAV73011	Aav73011 Human adu
7	684	33.7	1853	6	ABO92036	Abog92036 Human pol
8	675	33.2	4093	4	ABL14208	Ab114208 Drosophila
9	668.5	32.9	2304	2	AAQ66892	Aaq66892 Glycosylt
10	668	32.9	2615	2	AAQ003464	Aaq003464 CDNA enco
11	667.5	32.9	3923	2	AAQ003463	Aaq003463 CDNA enco
12	667.5	32.9	3950	13	ADT99580	Adt99580 Mouse bet
13	667.5	32.9	3953	1	ACN45213	Acn45213 Mouse MRN
14	666	32.8	1158	3	AAA53309	Aaa53309 Beta 1,4
15	666	32.8	1158	6	ABK90350	Abk90350 CDNA enco
16	666	32.8	1192	2	AAK80151	Aak80151 Human bet
17	666	32.8	1194	11	ADZ47902	Adz47902 Galactose
18	666	32.8	1197	10	ADDE68084	Adde68084 Human bet
19	666	32.8	1197	10	ADPF18447	Adpf18447 Human bet
20	666	32.8	1197	12	ADQ78357	Adq78357 Human bet
21	666	32.8	1265	2	ADQ31434	Adq31434 Encodes a
22	666	32.8	4072	12	ADQ87369	Adq87369 Human lum
23	666	32.8	4080	11	ADZ47903	Adz47903 Galactose
24	666	32.8	4162	10	ADJ56362	Adj56362 Human CDN
25	665.5	32.8	2304	2	AAQ66891	Aaq66891 Glycosylt
26	665	32.7	1265	2	AAQ66889	Aaq66889 Galactose
27	663.5	32.7	1116	13	ADQ91373	Adq91373 Manil-Gal
28	661	32.5	1119	10	ADJ94853	Adj94853 Novel NOV
29	661	32.5	1500	10	ADFI8449	Adfi8449 N-acetyl1g
30	660.5	32.5	1191	14	ADZ69854	Adz69854 Human gal
31	658	32.4	1116	14	ABE95896	Abe95896 Human bet
32	658	32.4	1119	10	ADPF73104	Adpf73104 Ebgr-1,4
33	658	32.4	1890	13	ADR20717	Adr20717 Human bet
34	658	32.4	1890	13	ACN37366	Acn37366 Tumour-as
35	658	32.4	2027	9	ACA63280	Ac63280 DNA encod
36	658	32.4	2027	10	AAD59475	Aad59475 Human bet
37	658	32.4	2027	12	ABE95864	Abe95864 Human bet
38	657.5	32.4	3194	12	ADQ24917	Adq24917 Human sof
39	656.5	32.3	905	13	ADQ80003	Adq80003 Bovine be
40	656.5	32.3	2163	12	ADQ78359	Adq78359 alpha-1,3
41	655.5	32.3	1068	10	ADDE68134	Adde68134 Human gal
42	655.5	32.3	1152	10	ADDE68140	Adde68140 Human gal
43	655.5	32.3	1152	10	ADDE68086	Adde68086 Human gal
44	655.5	32.3	6406	14	ADZ69841	Adz69841 Plasmid p
45	655	32.3	1197	13	ADQ79988	Adq79988 Human bet

ALIGNMENTS

RESULT 1		
ID	ADM32969	standard; cDNA, 1152 BP.
XX	ADM32969;	
AC	17-JUN-2004	(first entry)
XX		
DE	Nucleotide sequence of a beta1,4-N-acetylglactosaminyltransferase.	
XX		
KW	beta1; 4-N-acetylglactosaminyltransferase; beta4GALNCT; animal cell	
KW	glycoprotein; lactinac motif; LDN motif; GalNAcbeta4GlcNAc-R;	
KW	glycoprotein hormone; pituitary gland; glycoconjugate; parasite;	
KW	immune response; infection; enzyme; gene; ss.	
XX		
OS	Caenorhabditis elegans.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	1..1152
FT	/*tag= a	
FT	/product= "beta1,4-N-acetylglactosaminyltransferase"	
PN	WO2004024938-A2.	
PD	25-MAR-2004.	
XX		
PF	12-SEP-2003; 2003WO-US028833.	
XX		

PR 13-SEP-2002; 2002US-0411242P.
 XX
 PA (CDMM/) CUMMINGS R D.
 PA (KAWA/) KAWAR Z.
 PI Cummings RD, Kawar Z;
 XX
 XX WPI: 2004-270055/25.
 DR P-PSDB; ADM32968.
 XX
 XX New purified approximately 4 acetyl galactosaminyl transferase that is
 PT substantially free of other proteins, useful for transforming or
 PT transfecting host cells for producing substantially pure forms of the
 PT enzyme.
 XX
 XX Claim 5; Fig 1; 51pp; English.
 XX
 XX The present sequence encodes a beta1,4-N-acetyl galactosaminyl transferase,
 CC designated beta4GALNACT. The enzyme is required for the biosynthesis of
 CC animal cell glycoproteins. The enzyme functions to synthesise the
 CC lactinac or LDN motif GalNAc beta4GlcNAc-R. This motif is glycoprotein
 CC hormones produced by the pituitary gland, and is a major marker of
 CC glycoconjugates made by parasitic and non-parasitic invertebrates and may
 CC be implicated in host immune responses to infection. The protein and
 CC polynucleotides are useful for transforming or transfecting host cells
 CC for producing substantially pure forms of the enzyme, or in vitro, for
 CC formation of a LDN structure on proteins or peptides.
 XX
 XX Sequence 1152 BP; 314 A; 275 C; 259 G; 304 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No: 3.53e-223 Length: 1152
 Score: 2031.00 Matches: 383
 Percent Similarity: 100.00% Conserves: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 DB: 12
 US-10-661-430-1 (1-383) x ADM32969 (1-1152)
 QY 1 MetAlaPheArgHisLeuAlaValAlaArgLeuLysSerLeuLeuValLeuCyAlaVal 20
 DB 1 ATGGCTTTTGGCTATTTGGCAGTCGCGACGACTCAAGTCGTTGCTCGTACTTTGGCCGTT 60
 QY 21 LeuLeuLeuValHisAlaMetIleTyrLysIleProSerLeuTyrGluAsnLeuThrIle 40
 DB 61 CTCTATTAGTTCATGCAATGATTTATTAAGATTCATCGCTTACGAAACCTTACTATC 120
 QY 41 GlySerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla 60
 DB 121 GGCTCCCTCGACCTTATTTGCGAGTCGACGCAATGAGGAGGAGTCTCGGGAATATACGGCT 180
 QY 61 SerThrSerAspAspLeuLeuAspThrTTrpAsnSerThrPheSerProIleSerGluVal 80
 DB 181 TCCACTTCGAGATGATCTACTTGAATACGTAATTCACGTTTCCACGATTTCTGAAGTT 240
 QY 81 AsnGlnThrSerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeu 100
 DB 241 AATGAGACTGTTTATGAGAGACATTCGTCATTCCTGTTCCCGCAACCGACTCTT 300
 QY 101 GlnPheCyAsnGlnThrProHisLeuValGlyProIleArgValPheLeuAspGlu 120
 DB 301 CAATTCGTATATCGACACCTCCCACTGTCGGAACCAATCCGTATTCCTCGATGAG 360
 QY 121 ProAspPheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyGlnHisGlyMet 140
 DB 361 CCCGACTTCAAAACTCTCGAAGAAATCTATCCGACACGACGCGGTGACATGATG 420
 QY 141 ProLysAspCyValAlaArgHisArgValAlaIleLeuValProTyrArgAspArgGlu 160
 DB 421 CTTAAGGATTTGTGTGCAAGGCAATCGTGTCTATTTATGTGCGCTTAAGAGATCGTGA 480
 QY 161 AlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaLysGlnGlnLeuAsp 180

DB 481 GCACATTTGAGATATATGCTCCACATTTTGCATCGTTCCTGCCCAACACAAATGAGC 540
 QY 181 TyrAlaIlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMet 200
 DB 541 TATGCAATTTTATGTTGAGCAAGTGGGAAATCGAATCGAATTCGCGGAAACTTAATG 600
 QY 201 AsnValGlyTyrAspValAlaSerArgLeuTyrProTyrGlnCyPheIlePheHisAsp 220
 DB 601 AACGTTGGATACGACGTACATCACCCCTTACCCATGGCAATGCTTCACTTCAATGAT 660
 QY 221 ValAspLeuLeuProGluAspAspArgAsnLeuTyrThrCyProIleGlnProArgHis 240
 DB 661 GTCGATTTTACTCCCGCAAGATGACGTAACTGTACACGTGTCCATTCACACGCTCAT 720
 QY 241 MetSerValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyGly 260
 DB 721 ATGAGTGTAGCGCATTCATTAATTCATTTAAATCTTCATATTCGCGATCTTGGCGGA 780
 QY 261 IleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerAsnAspPheTyr 280
 DB 781 ATCAGTGCATTAACAAAGATCACCTGAGAAATCAATGATTTTTCGATATGATTTTTCG 840
 QY 281 GlyTyrGlyGlyGluAspAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysVal 300
 DB 841 GGTGGGCGCGAGAGGACGACGATTTGGCGACGAGAAATCATGATGGCTGAGTGAAGTT 900
 QY 301 SerArgTyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThr 320
 DB 901 TCAGATATTCGACACAAATTTGACGATTAATAATGATTAACGACTCGACGGAAGGAGAG 960
 QY 321 AsnProValAsnLysCyAspArgTyrLysIleMetGlyGlnThrLysArgArgTyrThrArg 340
 DB 961 AATCCAGTTAATTAAGCCCGCTTACAAATATATGCGCAACGAAAGCGCGATGAGACAGT 1020
 QY 341 AspGlyLeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThr 360
 DB 1021 GACGGCCTTAAGCAATCTGAATGATTAAGTCGTAAATCTGGAATTTGAAGCTCTCTCACT 1080
 QY 361 ArgAlaValValAspLeuLeuGluLysAspCyAspArgGluLeuArgArgAspPhePro 380
 DB 1081 CGAGCGGTGTCGATTTGCTCGAAAAGACTGCGCGGAGCTGCGAAGGACCTTCCCA 1140
 QY 381 ThrCyAspPhe 383
 DB 1141 ACCTGTTTT 1149
 DB
 RESULT 2
 AAC77836
 ID AAC77836 standard; cDNA; 2002 BP.
 XX
 AC AAC77836;
 XX
 XX 08-FEB-2001 (first entry)
 DT
 DE Human cancer associated gene sequence SEQ ID NO:230.
 XX
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
 KW antidiabetic; antisthmatic; antineumatic; antidiarrhetic; antiviral;
 KW antiinflammatory; antihypertoid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nocotropic;
 KW vasotropic; antiproliferative; angiogenesis; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; grief versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.
 XX
 OS Homo sapiens.
 XX
 XX NC0200055350-A1.
 PN
 XX
 PD 21-SEP-2000.

XX 08-MAR-2000; 2000MO-US005882.
 PF 12-MAR-1999; 99US-0124270P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Ruben SM;
 XX MPI, 2000-587533/55.
 XX P-PSDB; AAB43627.
 DR
 XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.
 PS Claim 1; Page 801-802; 2352pp; English.
 CC
 CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
 CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; vulnary; immunomodulator;
 CC antidiabetic; antiaesthetic; antineumatic; antirheumatic;
 CC antiinflammatory; antihypertoid; antiallergic; antibacterial; antiviral;
 CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
 CC neurotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 2002 BP; 428 A; 576 C; 527 G; 469 T; 0 U; 2 Other;

Alignment Scores:
 Pred. No.: 9.32e-72 Length: 2002
 Score: 718.00 Matches: 134
 Percent Similarity: 64.60% Conservative: 43
 Best Local Similarity: 48.91% Mismatches: 95
 Query Match: 35.35% Indels: 2
 DB: 3 Gaps: 2

US-10-661-430-1 (1-383) x AAC7836 (1-2002)

QY 95 ProaspaenGlnThleuGlnPheCysaenGlnThrProRohisIeuValIleProIle 114
 DB 470 CCARCTCCCTCAAGGCTCTGCTCTGCTCCAGAACGATCTCTCTGAGTGGTCTGTG 529
 QY 115 ArgVal----PheIeuasgluPProasprelysthrleuGlnIleIleYrProasphtr 133
 DB 530 TCGGTCCTTTAGCCCACTGCTGCAAGATGTGAGGAGGATCCCGGGTA 589
 QY 134 HisalagIyGlnIleGlyMetProIyAspCysValAlaArgHisArgValAlaIleIle 153
 DB 590 GAACGAGGGGGCGGATCGCGCTGAGGTTGAGCCCGCTCCGGAACACCATCATTT 649
 QY 154 ValProIyArgAspArgGluAlaHisIleuArgIleMetleuHisIleuHisIleu 173
 DB 650 GGGCTCATCTGCGCGGAGGACCACTGCTGCTCTCTACCACTGACCCCTTC 709
 QY 174 IeuAlaIyGlnIleuAspIyArgAlaIlePheIleValIleGluValAlaIleGlnIthr 193
 DB 710 TTGCAGCGCCGACGCTTCTTATGCACTTATGTCATTCACACGCTGGAATGGAACA 769
 QY 194 PheasArgIyIyIleuMetAsnValIyIyAspValAlaIleSerArgIleuYrProItr 213

DB 770 TTAAACAGGGCAAACTGTTGAACGTTGGGGTGGAGAGCCCTGCTGATGAAGATGG 829
 QY 214 GlnCysPheIlePheHisAspValAspIleuPProGluAspAspArgAsnIleuYrThr 233
 DB 830 GACTGCTGTTCTTTCACAGATGAGTGAATCTTGGCAGAAATGACACAAATCTGTATGTG 889
 QY 224 Cys---ProIleGlnProArgHisMetSerValAlaIleAspIyPheAsnIyIyIleu 252
 DB 890 TGTGACCCCGGGGAGCCCGGCAATGTTGCCGTTGTATGAACAAGTTTGGATACAGCCTC 949
 QY 253 ProIySerAlaIlePheGlyIyIleSerAlaIleuThrIyAspHisIleuIyIle 272
 DB 950 CCGTACCCCGGCACTTCTGAGAGTCTCAGCAGCTTACTCTCTGACACAGTACCTGAAAGTG 1009
 QY 273 AsnGlyPheSerAspAspPheIlePheGlyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 292
 DB 1010 AATGGCTTCCCAATGAATACTGGGGGTGGTGGATGAGACAGATTTGCTACAGG 1069
 QY 293 ThrSerMetAlaGlyIleuIyValSerArgIyYrProThrGlnIleAlaArgIyIyIyMet 312
 DB 1070 GTGGCCCTGGCTGGATGAATCTCTGGCCCGGCACATGCTGATGACACTATGAAGTG 1129
 QY 313 IleIyHisIserThrGluAlaThrAsnProValAsnIyCysArgIyIyIleMetGly 332
 DB 1130 GTGAAGCACCGAGAGATGAAGGCAATGAGGAAATCCCGACAGTTTGACCTCTGCTC 1189
 QY 333 GlnThrIyAspArgIyThrIyArgAspGlyIleuSerAsnIleuIyIyIyIyIyIyIyIyIy 352
 DB 1190 CGTACCCGAAATTCCTGAGCAGAGATGGATGAATCTGACATGACATACAGTTGCTGCT 1249
 QY 353 IeuGlnIleuIyProIeIyIyThrArgAlaValAlaIleu 366
 DB 1250 CGAGAGCTGGGGCTCTTTATACCAACATCACAGAGACATT 1291
 RESULT 3
 ID AA242120 standard; cDNA, 1938 BP.
 AC AA242120;
 XX 31-JAN-2000 (first entry)
 DT
 XX Human endometrium tumour cDNA derived EST 140.
 DE
 XX Endometrium; human; tumour; cancer; anticancer; cytostatic;
 KW EST: treatment; uterine; gene therapy; expressed sequence tag; ss.
 XX Homo sapiens.
 OS
 XX DE19817948-A1.
 PN
 XX 21-OCT-1999.
 PD
 XX 17-APR-1998; 98DE-01017948.
 PF
 XX 17-APR-1998; 98DE-01017948.
 PR
 XX (META-) METAGEN GBS GENOMFORSCHUNG MBH.
 PA
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 PI MPI, 1999-591957/51.
 DR
 XX New nucleic acid sequences expressed in uterine cancer tissues, and
 PT derived polypeptides, for treatment of uterine and endometrial cancer and
 PT identification of therapeutic agents.
 PS Claim 3; Page 273; 444pp; German.
 CC This invention describes novel human nucleic acid (cDNA) sequences (A),
 CC that are highly expressed in uterine tumour tissue and which have
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)

CC are used (i) to identify agents suitable for treatment of uterine or
CC endometrial cancer; (ii) directly for treating these forms of cancer
CC (including expression from gene therapy vectors) and (iii) for generation
CC of specific antibodies. (A) are identified by assembling ESTs (expressed
CC sequence tags) from a particular tissue type before comparison of
CC expression patterns. This allows a significantly longer fragment of the
CC gene to be revealed, so should reduce the number of failures associated
CC with the fact that ESTs from different libraries may represent different
CC parts of the same unknown gene, distorting the estimated frequency of
CC occurrence in a particular tissue. AA241981-242121 represent EST
CC fragments derived from a human endometrium tumour cDNA library which
CC encode the protein sequences represented in AAY59941-Y60328
XX

XX Sequence 1938 BP; 395 A; 567 C; 515 G; 461 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.16e-71	Length:	1938
Score:	717.00	Matches:	134
Percent Similarity:	64.60%	Conservative:	43
Best Local Similarity:	48.91%	Mismatches:	95
Query Match:	35.30%	Indels:	2
DB:	2	Gaps:	2

US-10-661-430-1 (1-383) x AA242120 (1-1938)

```
Qy 95 ProAspAsnGlnThrLeuGlnPheCysAsnGlnThrProHisLeuValGlyProIle 114
Db 432 CCAAGCTCCTCAAGGTCGCCCTACTGTCAGAAAGATCTCTCTTAAAGGCTCCTG 491
Qy 115 ArgVal--PheLeuAspGluProAspPheLeuThrLeuGluValIleTyProAspThr 133
Db 492 TCGGTGCTCTTAGCCAGTGCATCTGCGAGAGATTGTGGCGAGATCCCGGGTA 551
Qy 134 HisAlaGlyGlnHisGlyMetProLysAspCysValAlaArgHisArgValAlaIleIle 153
Db 552 GAACCAAGGAGCGGTCACCCCTGCGAGGTTGTGAGCCCGCTCCCAACAGCATCATTT 611
Qy 154 ValProTyArgAspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeu 173
Db 612 GTGCCTCATGTCGCGCGGAGACACCACTGCTGCTCTACCACTGCACCCCTTC 671
Qy 174 LeuAlaLysGlnLeuAspTyArgAlaIlePheIleValGluGlnValAlaAsnGlnThr 193
Db 672 TTGCAGCGCAGAGCTTGTGATGATCATCTCATCCACCGCTGGAATGGAACA 731
Qy 194 PheAsnArgGlyLysLeuMetAsnValGlyTyTrpAspValAlaSerArgLeuTyProTrp 213
Db 732 TTTAACAGGCGCAAACTGTGAACGTTGGGTCGAGAGGCTGCGTGAATGAAGTGG 791
Qy 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyThr 233
Db 792 GACTGCTGTTTTCACAGATGTGACCTCTTGCACAATAATGACCAATCTGTATG 851
Qy 234 Cys--ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyLysLeu 252
Db 852 TGTGACCCCGGAGACCCCGCATGTGTCCTGCTGAACAAGTTGGATGACAGCTTC 911
Qy 253 ProTyTrpSerAlaIlePheGlyIleSerAlaLeuThrLysAspHisLeuLysIle 272
Db 912 CCGTACCCCACTGACTTGGAGAGTCTCAACACTTACTCTGACCAAGTCTGAAGATG 971
Qy 273 AsnGlyPheSerAsnAspPheTyArgIleTyProGlyGlnAspAspAspLeuAlaThrArg 292
Db 972 AATGCTTCCCATGATGATCTGGGCTGGGTGTGATGATGACGATCTGATCAAG 1031
Qy 293 ThrSerMetAlaGlyLeuLysValSerArgTyProThrGlnIleAlaArgTyLysMet 312
Db 1032 GTGCGCTGCTGGATGATGATCTCGGCCCCCACTGTAGAGACACTATAAGATG 1091
Qy 313 IleuHisSerThrGlnAlaThrAsnProValAsnLysCysArgTyLysIleMetGly 332
Db 1092 GTGAAGACCCGAGAGATGAAGGCAATGAGAAATCCCAAGATTGACTCTCGTCTC 1151
```

```
Qy 333 GlnThrLysArgArgTyThrArgAspGlyLeuSerAsnLeuLysTyLysLeuValAsn 352
Db 1152 CGTACCCCAATTCCTCGACGCAAGATGGAGATCACTGATCAACGATTTGCTGCT 1211
Qy 353 LeuGlnLeuAspProLeuTyThrArgAlaValAlaAspLeu 366
Db 1212 CGAGACTGCGGGCTCTTATATCAACATCAACAGACGACATT 1253
```

RESULT 4

ID ACN37523 standard; cDNA, 1949 BP.

ACN37523;

18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) cDNA DNA323958, SEQ ID NO:468.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KM tumour; diagnosis; cell proliferative disorder; breast cancer;
KM colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KM central nervous system cancer; bladder cancer; pancreatic cancer;
KM cervical cancer; melanoma; leukaemia; hybridisation probe;
KM chromosome identification; chromosome mapping; gene mapping;
KM gene therapy; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX 29-SEP-2003; 2003WO-US028547.

XX 02-OCT-2002; 2002US-0414971P.

XX (GENTH) GENENTECH INC.

XX Wu Td, Zhang Z, Zhou Y;

XX WPI; 2004-347921/32.

XX P-P-SDB; ABW80189.

PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.

XX Claim 1; SEQ ID NO 468; 7273bp; English.

XX The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to the TAT nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acids and
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods. TAT polypeptides, nucleic acids,
XX antibodies of cancer in mammals. TAT polypeptides, nucleic acids,
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukaemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention
XX
XX Sequence 1949 BP; 376 A; 584 C; 525 G; 464 T; 0 U; 0 Other;


```

QY 113 Pro11eArgValPheLeuAerGluProAerPheLeuGlu--Lys11eTyPro 131
DB 535 CCATACACGCCCAACAGACAGACAGTGCATCTGACCTTTATGAGGGAGGAGCTTGACCT 594
QY 132 AepThiHia1aGlyGlyH1sGlyMetProLysAerCyValA1AaRghiSarValA1a 151
DB 595 CTTTTCGCCCTGCGGCTGCGCTTGCAGCTGAAGAAATGCAATGCCAGCTACGCTGCT 654
QY 152 Ile11eValProTyraArgAerArgGluA1aH1sLeuAerG11eMetLeuH1sAenLeuH1s 171
DB 655 ATGTGTGTGCTTGCAGTGCATACAGCCCATCTTATCTTCTTCCGCGCAATCTCAC 714
QY 172 SerLeuLeuA1aLysGlnGlnLeuAerTyra11ePheH1eValGlnGlnValA1aAen 191
DB 715 CCATTTCTGTAGAGAGCGGCATCGCTATCGCATTTTCTATTGTAGAGAGAGCAACCGG 774
QY 192 GlnThrPheAerAerGlyLysLeuMetAerValA1GlyTyraValA1aSerAerGleuTy 211
DB 775 AAGCCCTTATCGGCTGCGCTGATGACATGTGTTATTTGAGAGCGCTTAAAGCTGTAC 834
QY 212 ProTyraGlnCyPheH1ePheH1sAerValA1aSerLeuProGluAerAerAerLeu 231
DB 835 CAGTGGGATTTTATTTATTTCCAGATGTCATCTTCTTGCCTTGAAGAGCGCAATCTT 894
QY 232 TyrThrCyPhePro11eGlnProAerGhiMetSerValA1a11eAerLysPheAerTy 251
DB 895 TACAACGTGTCAAGTGCAGCGGCAGACATGTCAGTGCCTTATAGACAGCGCAATCTTACG 954
QY 252 LeuProTyraSerA1a11ePheGlyLys11eSerA1aLeuThryAerH1sLeuLys 271
DB 955 TTCCCTTATCGATCAATATTTGGAGGTGTTCCGCAATGACGCGTGAAGCTTTCAGGCC 1014
QY 272 11aAerGlyPheSerAerAerPheThryProGlyLysGlnAerAerAerAerA1aThr 291
DB 1015 G1AAATGGGTCTCAAACTGTTTGGCGGGCGGCGAGATGAGACATGTCAC 1074
QY 292 ArgThrSerMetA1aGlyLeuLysValSerAerTyraProThryGln11eA1aerTy 311
DB 1075 AGGTGGAAGACGCAAACTTATTCATTCAGATGTCGGTCAACATAGCTCGGTACAG 1134
QY 312 Met11eLysH1sSerThryGluA1aThrAerProValA1aLysCyAerTyraLys11eMet 331
DB 1135 ATGCTGAAGCATCAGAAAGAAAGCCAACTCTAAG-----CCCTTGAAGAACTTA 1185
QY 332 GlyGlnThryAerAerTyraProThryAerAerGlyLeuSerAerAerAerAerA1a 351
DB 1186 CAGAAATGGCATGACCAAAATTAAGAAAGAGATGAAATCACTGATTAAGATTTCTATCTAC 1245
QY 352 AseLeuGluLeuLysProLeuTyraThryA1aValA1aSerLeu 366
DB 1246 AGCATCAAAACAATTTCCAACTTTCAGTACTTACAGAGCTA 1290
RESULT 6
ID AAV73011 standard; CDNA; 1853 BP.
XX
AC AAV73011;
XX
DT 15-MAR-1999 (first entry)
XX
DE Human adult brain secreted protein ges1_1 cDNA.
XX
KM Secreted protein; protein factor; human; ges1_1; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 154..975
FT /tag= a
FT /note= "polynucleotides comprising nucleotides 154-972
FT and 1-341 are also claimed"
XX
PN WO9842741-A2.

```

```

XX
PD 01-OCT-1998.
XX
PF 25-MAR-1998; 98WO-US005972.
XX
PR 25-MAR-1997; 97US-00825145.
PR 24-MAR-1998; 98US-00046881.
XX
PA (GEM ) GENETICS INST INC.
XX
PI Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MO;
DR P-PSDB; AAM82006.
XX
PT New polynucleotides encoding secreted human proteins - are derived from
PT human foetal brain, adult testes, adult brain, or placenta cDNA
PT libraries, potentially useful as, e.g. immunostimulants or
PT antiinflammatory agents.
XX
PS Claim 35a; Page 85-86; 113pp; English.
XX
CC This cDNA sequence codes for novel human secreted protein ges1_1 (see
CC AAM82006). The cDNA was isolated from a human adult brain cDNA library
CC using methods which are selective for cDNAs encoding secreted proteins,
CC or was identified as encoding a secreted or transmembrane protein on the
CC basis of computer analysis of the amino acid sequence of the encoded
CC protein. The nucleotide sequence shows homology to some known database
CC sequences. ges1_1 is a full-length clone and can be obtained from the
CC deposit clone ATCC 98371 as a 1850 bp EcoRI/NcoI fragment. The invention
CC provides polynucleotides (see AAV73000-12), derived from human foetal
CC brain, adult testis, adult brain and adult placenta cDNA libraries, that
CC encode novel secreted proteins (see AAM81997-07). These can be isolated
CC from composite clone ATCC 98371 using specific probes (see AAV73013-22).
CC The polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals (no supporting data
CC are given). Suggested activities include nutritional, immune stimulating
CC (e.g. as vaccines) or suppressing, haematopoietic regulating, tissue
CC growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, antiinflammatory, cachectin/tumour invasion
CC suppressor, and tumour inhibition activities. The polynucleotides are
CC also stated to be useful for gene therapy
XX
SQ Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;
XX
Alignment Scores:
Pred. No.: 6,79e-68 Length: 1853
Score: 684.00 Matches: 133
Percent Similarity: 63.87% Conservative: 42
Best Local Similarity: 48.54% Mismatches: 97
Query Match: 33.68% Indels: 3
DB: 2 Gaps: 2
XX
US-10-661-430-1 (1-383) x AAV73011 (1-1853)
QY 95 ProAerAerGlnThryLeuGlnPheCyAerGlnThryProProH1sLeuValGlyPro11e 114
DB 358 CCAAGCTCTCAAGAGTGTGCGCTGCTCCAGAAAGATCTCTCTTATGAGGGATCCGTG 417
QY 115 ArgVal---PheLeuAerGluProAerPheTyraThryGlnGlyLys11eTyraProAerTh 133
DB 418 TCGGTGCTTTAGCCCAAGTGCATGACAGATGAGATTTGTGAGGAGATCCCGGATA 477
QY 134 H1a1aGlyGlyH1sGlyMetProLysAerCyValA1aAerGhiSarValA1a11e 153
DB 478 GAACACAGGGGGCGGATACCGCTGCAAGTTGAGAGCCCGCTCCGAAACAGCATCAT 537
QY 154 ValProTyraArgAerArgGluA1aH1sLeuAerG11eMetLeuH1sAenLeuH1sSerLeu 173
DB 538 GTGCTTCATGTGCGGAGACACACATGCGCTGCTCTTACCACTGACACCCCTTC 597

```

```

Oy 174 LeuAlaIySGInGInLeuAspTyrAlaIlePheIleValGInValAlaAsnGInThr 193
Db 598 TTGACGCCCGACAGCTTCTTATGGCATCTATGATCCACAGGCTGGAAATGAAACA 657
Oy 194 PheAsnArgIyLyLeuMeaAsnValGlyTyrAspValAlaSerArgLeuTyrProTrr 213
Db 658 TTTAAACAGGCGAAACCTGTGAACCTGGGGTGCAGAGGCCCTCGTATGAAAGATGG 717
Oy 214 GInCysPheIlePheIleAspValAspLeuProGInuAspArgAsnLeuTyrThr 233
Db 718 GACTGCTCTTCTTGCAGATGTGACCTCTTGCAGAAAGACACACATCTGTATGTG 777
Oy 234 Cys---ProIleGInProArgHisMetSerValAlaIleAspLyPheAsnTyrLyseu 252
Db 778 TGTGACCCCGGGAGACCCCGCATGTGCGCTGTGTAAGAAATTTGGATACAGCTTC 837
Oy 253 ProTyrSerAlaIlePheGlyIyIleSerAlaLeuThrLyAspHisLeuLyAsIle 272
Db 838 CCGTACCCCGACAGTACTTGCAGAGAGTCTCAGACATTACTCTGACAGTACTGAAGATG 897
Oy 273 AsnGlyPheSerAsnAspPheTrrGlyTrrGlyGlyuAspAspAspLeuAlaThrArg 292
Db 898 AATGCTTCCCAAGAAATACCTGGGCTGGGGTGGATGAGACACTT-GCTACCAAGG 956
Oy 293 ThrSerMetAlaGlyLeuLyAsValSerArgTyrProThrGInIleAlaArgTyrLyMet 312
Db 957 GTGCGCCCTGGCTGGAGATGAATCTCTGGCCCCCACAATCTGTAGACACTATAGATG 1016
Oy 313 IleLyHisSerThrGlnAlaThrAsnProValAsnLyCysArgTyrLyIleMetGly 332
Db 1017 GTGAAGACACCGAGAGATTAAGGGCAATGAGAAATCCCAAGATTGACTCTCTGGTTC 1076
Oy 333 GInThrLyAsArgTyrThrArgAspGlyLeuSerAsnLeuLyTyrLyLeuValAsn 352
Db 1077 GCTACCCAGAAATCTCTGACGAGATGGAGTGAACCTCAGACATACAGTTGCTGCT 1136
Oy 353 LeuGlyLeuLyProLeuTyrThrArgAlaValAlaAspLeu 366
Db 1137 CGAGAGCTGGGGCTCTTTATTCACAACTACAGACAGACATT 1178

```

RESULT 7
ABQ92036
ID ABQ92036 standard; cDNA; 1853 BP.
XX
XX ABQ92036;
AC
XX
XX 04-OCT-2002 (first entry)
DE Human polynucleotide SEQ ID NO 33.
XX
XX Human; cytostatic; antirheumatic; antiarthritic; vulnery; analgesic;
XX antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
XX neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
XX anticancer; fungicide; antidiabetic; antiallergic;
XX immunostimulant; antiparasitic; secreted protein; transmembrane protein;
XX cytokine; cell proliferation; cell differentiation; autoimmune disease;
XX stem cell; growth factor; nervous system disease; neuropathy;
XX Alzheimer's disease; Parkinson's disease; Huntington's disease;
XX osteoporosis; severe combined immunodeficiency; SCID; infection;
XX multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
OS Homo sapiens.
PN US2002065394-A1.
PD 30-MAY-2002.
XX
XX 22-DEC-2000; 2000US-00745763.
PF
XX 18-MAR-1998; 98US-00040963.
PR
XX
XX (JACOBI) JACOBS K.
PA (MCCOY) MCCOY J M.

```

PA (LAWA/) LAWALLIE E R.
PA (COLL/) COLLINS-RACIE L A.
PA (EVAN/) EVANS C.
PA (MERB/) MERBERG D.
PA (TREA/) TREACY M.
PA (SPAUV/) SPAULDING V.
PI Jacobs K, McCoy JM, Lawallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Spaulding V;
XX
XX WPI: 2002-582343/62.
XX P-PSDB; ABP61819.
XX
XX Novel secreted or transmembrane protein and polynucleotide encoding the
XX protein, useful for diagnosis and treatment of neurological disorders,
XX cancer, autoimmune disease, bone disorders and lung or liver fibrosis.
XX
XX Claim 111, Page 144-145; 284pp; English.
XX
XX The invention relates to human secreted or transmembrane protein (I),
XX their fragments and is encoded by specific complementary deoxyribonucleic
XX acid (cDNA) inserts (II), where the protein is substantially free from
XX other mammalian proteins. (I) are useful for preventing, treating or
XX ameliorating a medical condition, especially immunological treatment or
XX prevention of tumours. (I) exhibits activity relating to angiogenesis,
XX cytokine, cell proliferation, cell differentiation, antiinflammatory,
XX stem cell growth factor activity and activin or inhibin-related
XX activities. (I) can be used to manipulate stem cells in culture to give
XX rise to neuroepithelial cells that can be used to augment or replace
XX cells damaged by illness, autoimmune disease, accidental damage or
XX genetic disorders. (I) induces the proliferation of neural cells and
XX regeneration of nerve and brain tissue and is useful for the treatment of
XX central and peripheral nervous system diseases and neuropathies, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis. (I) is involved in chemotactic or chemokinetic
XX activity, regulation of haematopoiesis and is useful for treating myeloid
XX or lymphoid cell disorders, platelet disorders such as thrombocytopaenia
XX and for regeneration of bone, cartilage, tendon, ligament and/or nerve
XX tissue growth and in tissue repair, healing of burns, incisions, ulcers,
XX for treating osteoporosis, osteoarthritis, bone degenerative disorders or
XX regeneration and treatment of lung or liver fibrosis, reperfusion injury
XX in various tissues, various immune deficiencies and disorders including
XX severe combined immunodeficiency (SCID), bacterial or fungal infections,
XX autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
XX diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
XX such as asthma or other respiratory problems. (II) is useful to express
XX recombinant protein, as markers for tissues in which the corresponding
XX CC protein is preferentially expressed and in gene therapy. The present
XX sequence is that of a polynucleotide of the invention
XX
XX SQ Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;
XX
XX Alignment Scores:
XX Pred. No.: 6,79e-68 Length: 1853
XX Score: 684.00 Matches: 133
XX Percent Similarity: 63.87% Conservative: 42
XX Best Local Similarity: 48.54% Mismatches: 97
XX Query Match: 33.68% Indels: 3
XX DB: 6 Gaps: 2
XX
XX US-10-661-430-1 (1-383) x ABQ92036 (1-1853)
Oy 95 ProAspAsnGInThrLeuGlnPheCysAsnGInThrProGInuAspValGlyProIle 114
Db 358 CCAAGCTCTCAAGAGTCCCTCTATGTCAGAAAGATCTCTCTTGTGGTCTGTG 417
Oy 115 ArgVal---PheLeuAspGluProAspPheLyThrLeuGInuLyIleTyrProAspThr 133
Db 418 TCGGTCCTTTAGCCCAAGTGCATCTGCGAGATGTTGGACGGAAATCCCGGGTA 477
Oy 134 HisAlaGlyGInuHisGlyMetProLyAspCysValAlaArgHisArgValAlaIleIle 153

```

```
Db 478 GAACCAAGGGGCGGACCGCTGACGATTGTGAGCCCGCTCCGGAACGACCATCAT 537
Qy 154 ValProTyrArgAspArgGluAlaHisLeuArgILeMetLeuHisBAenLeuHisSerLeu 173
Db 538 GAGCCCTATGCTGCGGAGACACCACTGCGCTGCTTAACCACTGACACCCCTTC 557
Qy 174 LeuAlaLysGlnGlnLeuAspTyrAlaIlePheIleValGluGlnValAlaAenGlnThr 193
Db 598 TTGCACGCGCAGACGCTTGTGTTATGGCATCTATGTATCATCAACAGGTGTAATGAAACA 657
Qy 194 PheAsnArgGlyLysLeuMetAsnValGlyTyrAspValAlaSerArgLeuTyrProTyr 213
Db 658 TTAAACAGGGCAAAACGTGTGAACGTTGGGGTGCAGAGGCGCTGCTGATGAAGAGTGG 717
Qy 214 GlnCysPheIlePheHisAspValAspLeuLeuProGluAspAspArgAsnLeuTyrThr 233
Db 718 GACTGCTGTTGTTGACGATGTGACCTTGTGCAAAATATGACCAATCTGATGTG 777
Qy 234 Cys---ProIleGlnProArgHisMetSerValAlaIleAspLysPheAsnTyrLysLeu 252
Db 778 TGTGACCCCGGGGAGCCCGCCCATGTTGCGGTTGCTATGAACAAGTTTGATCAGCCTC 837
Qy 253 ProTyrSerAlaIlePheGlyLysIleSerAlaLeuThrLysAspHisLeuLysIle 272
Db 838 CCGTACCCCGCAGTACTCGAAGAGTCTCAGCACTTACTCTGACCACTGACCTGAAGATG 897
Qy 273 AsnGlyPheSerAspAspPheTyrGlyTyrGlyGlyValAspAspAspLeuAlaThrArg 292
Db 898 AATGGCTTCCCAATGATATCTGGGGCTGGGGTGGTGGATGAGATGACCTT-GCTACCAAG 956
Qy 293 ThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyrLysMet 312
Db 957 GTGCGCTGCTGGATGAGATCTCTCGGCCCCCACTCTGATGAGACACTAAGATG 1016
Qy 313 IleLysHisSerThrGluAlaThrAspProValAsnLysCysArgTyrLysIleMetGly 332
Db 1017 GTGAAGACACGAGAGATTAAGGGCATGAAGAAATCCCAAGATTGTGACCTTCGTGTC 1076
Qy 333 GlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLysTyrLysLeuValAsn 352
Db 1077 CGTACCCAGAAATCTCTGACGCAAGATGGATGAATCACTGACATATCCAGTTGCTGCT 1136
Qy 353 LeuGlnLeuLysProLeuTyrThrArgAlaValAlaValAspLeu 366
Db 1137 CGAGAGCTGGGGCTCTTTATACCAATCAACAGACACATTT 1178

RESULT 8
ABLI4208/c
ID ABLI4208 standard; cDNA; 4093 BP.
XX
XX ABLI4208;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 37106.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX MO200171042-A2.
XX
XX PD 27-SEP-2001.
XX
XX PF 23-MAR-2001; 2001WO-US009231.
XX
XX PR 23-MAR-2000; 2000US-0191637P.
XX
XX PR 11-JUL-2000; 2000US-00614150.
XX
XX PA (PEKE ) PE CORP NY.
XX
XX PI Venter JC, Adams M, Li FWD, Myers EW;
```

```
XX
XX WPI; 2001-656860/75.
DR P-PsDB; ABB70105.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 37106; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 4093 BP; 1129 A; 905 C; 927 G; 1132 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,38e-66 Length: 4093
Score: 675.00 Matches: 161
Percent Similarity: 50.00% Conservative: 59
Best Local Similarity: 36.59% Mismatches: 137
Query Match: 33.23% Indels: 83
DB: Gaps: 11
US-10-661-430-1 (1-383) x ABLI4208 (1-4093)
Qy 5 HisLeuAlaValAlaArgLeuLysSerLeuLeuValLeuCysAlaValLeuLeuVal 24
Db 2852 CATATGCTCTGGCAGAGGATATCTTATCAATATTGTTGGACGAGATTTGCACTT-- 2795
Qy 24 HisAlaMetIleTyrLysIleProSerLeuTyrGluAsnLeuThr-----IleG1 41
Db 2794 ----CGGCGGTGCACAGATAGTGCATATCTACGAGAACGTAAGCAGATGAGACCGG 2739
Qy 41 YSerSerThrLeuIleAlaAspValAspAlaMetGluAlaValLeuGlyAsnThrLase 61
Db 2738 AGCGACGTGA--GCATCCAGGCTGCCCTCCCGCTCGCTTCAAAACACAGAGA 2682
Qy 61 rThrSerAspAspLeuAspThrTrpAsnSerThrPhe---SerProIleSerGluVal 80
Db 2681 GCGGACACAGAGCTAATAGCGGACCCACATCTTACATTAAGAACTGTGATGCCACGGC 2622
Qy 80 LAsnGlnThrSerPheMetGluAspIleArgProIleLeu-----Ph 94
Db 2621 AAATCTTACTTTCATTCACAAAGATTAAGCGGCTTCGTGCGTGGCAAAAGAAATTTT 2562
Qy 94 eProAspAsnGln-----ThrLeuGlnPheCysAsnGln----- 105
Db 2561 GCCCGCGCAGAAATCCACATCCGCTCTTGGCACTGCATGCATCCGATCCGATCCCG 2502
Qy 106 -----ThrProPheIle 110
Db 2501 TGAATGTAAGCTGGCTGCTGCTTTTATGACAGATTTAAATACGTTCACTGTTT 2442
Qy 110 uValGlyProIleArgValPheLeuAspGluProAspPheLysThrLeuGlu--LysI1 129
Db 2441 AGGTGAGCCCATACCGCCCAACAGACGACGTGACGTGATGATGAGCGGAGCT 2382
Qy 129 eTyrProAspThrHisAlaGlyLysIleGlyMetProLysAspCysValAlaArgHisAr 149
Db 2381 TGAACCTTTTGGCGCTGGTGGCGCTTGCAGTGAAGAACTGCATGCAATGCCAGATCA 2322
Qy 149 gValAlaIleIleValProGlyTrpArgAspArgGluAlaHisLeuArgIleMetLeuHisAs 169
Db 2321 CGTGGCTATGTGTGCTTCCGCGATGATAGCGCCATCTATTACTTTCTTGGCGCAA 2262
```

```
QY 169 nleuHisSerLeuLeuAlaValGlnGlnLeuAspTyrAlaIlePheIleValGlnGlnVala 189
DB 2261 CATCCACCCATTCTTGATGATGAGCAGCGCATCCGCTATGCGATTTCATTGATGAGAGAGC 2202
QY 189 lAlaAsnGlnThrPheAsnArgGlyValLeuMetAsnValGlyTyrAspValAlaSerArg 209
DB 2201 CAACGGGAAGCCCTTAAATCGGGCTGCCATGATGAACATGTTGTTATTTGGAGGCCCTTAA 2142
QY 209 GHeuTyrProTTPGlnCysPheIlePheHisAspValAspLeuLeuProGlnAspAspArg 229
DB 2141 GCTGTAACAGTGGGATGTTTATATATCCACATGTCGATCTTCGCTTGGACGACCG 2082
QY 229 GAsnLeuTyrThrCysProIleGlnProArgHisMetSerValAlaIleAspLysPheAs 249
DB 2081 CATCTTTTAACTGTCACGTCAGCCGCGACACATGTCAGTGGCTATGACACCGTGA 2022
QY 249 n-----Ty 250
DB 2021 CTTTCAGCAGATGAGCAAGAAGATTAACAATCAATGTAATGATTCAATCTTTA 1962
QY 250 rlyLeuProTyrSerAlaIlePheGlyGlyIleSerAlaLeuThrLysAspHisLeuLys 270
DB 1961 TAGGTTCCTTATCGATCAATATTTGGAGGTGTTCGCAATGACGCGGAGCCACTTTCA 1902
QY 270 slyAlleAsnGlyPheSerAsnAspPheTyrGlyTyrGlyGlyLysAspAspLeuAl 290
DB 1901 GGGCGTAATGGGCTTCTCAAACTCGTTCTTGCGTGGCGCGGAGATGACGATGTC 1842
QY 290 aThrArgThrSerMetAlaGlyLeuLysValSerArgTyrProThrGlnIleAlaArgTyr 310
DB 1841 CAACAGGTGGAAGCAGCAGCAACCTATTCATATCAAGGTATCCGGTCACATAGCTCGCTA 1782
QY 310 rlyMetIleLysHisSerThrGluAlaThrAsnProValAsnLysCysArgTyrLysI 330
DB 1781 CAGAGTCTGAAGCATCAGAGGAAAGGCGCAATCTTAACGGCTAGTAGACCATATCTAT 1722
QY 330 eMet----- 331
DB 1721 TATCCATATTTAGACTAACAATATGAAATTTTACCTTCTTTTGGACAGCTATGA 1662
QY 332 -----GlyGlnThrLysArgArgTyrThrArgAspGlyLeuSerAsnLeuLys 347
DB 1661 AACTTACAGAAATGGCATGAGCAAAATAGAA-----CAGGATGGATCAACTCGATAAA 1608
QY 347 sTyrLysLeuValAsnLeuGlnLeuLysProLeuTyrThrArgAlaValAlaAspLeu 366
DB 1607 GTATTCTATCTACAGCATCAAAACATTTTCAACTTTCGATCTTAGCAGAGCTA 1550
RESULT 9
AA066892
ID AA066892 standard; cDNA; 2304 BP.
XX
AC AA066892;
XX
DT 25-MAR-2003 (revised)
DT 18-NOV-1994 (first entry)
XX
XX Glycosyltransferase hybrid.
XX
KW Galactosyltransferase; sialyltransferase; hybrid protein;
KW glycosyltransferase; glycoprotein; glycolipid; oligosaccharide; YEPGSTB;
KW glycosylation; Saccharomyces cerevisiae; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2304
FT /tag= a
FT /product= "galactosyltransferase-sialyltransferase hybrid
FT protein."
XX
XX MO9412646-A1.
```

```
PD 09-JUN-1994.
XX
XX 15-NOV-1993; 93WO-BP003194.
XX
XX 27-NOV-1992; 92EP-00810924.
XX
XX (CIBA ) CIBA GEIGY AG.
XX
PI Berger EG, Matzele M, Iwanow SX;
XX
XX WPI, 1994-200274/24.
XX
XX P-PSDB; AAR55709.
XX
PT Proteins with glycosyl transferase activity - useful for synthesis or
XX modification of glyco-proteins, glyco-lipid(s) and oligosaccharide(s).
XX
XX Disclosure; Page 52-57; 67pp; English.
XX
XX PS
XX
XX Hybrid glycosyltransferases (see also AAR55708) consisting of a membrane-
XX bound galactosyltransferase linked at its C-terminal to the N-terminal of
XX a soluble sialyltransferase have been expressed in Saccharomyces
XX cerevisiae using plasmid YEPGSTA. (Updated on 25-MAR-2003 to correct PN
XX field.)
XX
SQ Sequence 2304 BP; 552 A; 658 C; 576 G; 518 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 5.69e-66 Length: 2304
Score: 668.50 Matches: 140
Percent Similarity: 54.18% Conservative: 74
Best Local Similarity: 35.44% Mismatches: 132
Query Match: 32.91% Indels: 49
DB: Gaps: 6
US-10-661-430-1 (1-383) x AA066892 (1-2304)
QY 14 LeuLeuValLeuCysValAlaValLeuLeuValHisAlaMetIleTyr----- 29
DB 76 CTCGTGGCGCGTGTGGCTCTGTGACCTTGGCGTACCCGCTTACTACTGCTGGCGGC 135
QY 30 -----LysIleProSerLeuTyrGlnAsnLeuThrIleGlySerSerThrLeuIle 46
DB 136 GACGTGAGCCGCGTCCCAACTG-----GTGCGAGTCTCCACACCGCTG 180
QY 47 AlaAspValAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
DB 181 CAGGCGGCTCGAACAAGTCCGCGCCGATCGGAGTCTCGGGAGACTCCGACCGGA 240
QY 61 -----SerThrSer 63
DB 241 GGGGCGCGCGCGCGCTCTAGGCGGCTCTCCAGCGCGCGCGGCGGCGGCGCATTC 300
QY 64 AspAspLeuLeuAspThrThrAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
DB 301 AGCCGAGTGTGATTTCTGGCGCTGGCGCGCTAGCACTTGAACCTCGGTC----- 351
QY 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
DB 352 -----CCAGTGGCCCAACACACCGGCACTGTGCGTGGCGCGCTGC 390
QY 104 AsnGlnThrProProHisIleLeuValGlyProIleArgValPheLeuAspGluPro--Asp 122
DB 391 CCTGAGAGTCCCGCGCTGTGGGCGCCATGCGATGAGTTGAATTAACATGCTGTGAGC 450
QY 123 PheLysThrLeuGluLysIleTyrProAspThrHisAlaGlyGlyHisGlyMetProLys 142
DB 451 CTGAGCTCTGTGGCAAAAGCAGAACCAATGGAAGATGGGCGCGCTATGCGCCAGG 510
QY 143 AspCysValAlaAsnHisArgValAlaIleIleValAlaProTyrArgAspArgGluAlaHis 162
DB 511 GACTGCGTCTCTCTCACAGAGTGGCAATCATTCATTCATTCGCAACGGGAGGAGCAGC 570
QY 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuAlaValGlnGlnLeuAspTyrAla 182
```

```

Db      571 CTCAGACTAGGCTATATTTTGGACCCAGCTCTGAGCGGCAGGAGCTGACTATGGC 630
Qy      183 ILePheIleValIgluInValIalaenGIInThrPheAsnArgIlyIyleuMetAenVal 202
Db      631 ATCTATGTTATCAACGAGCGGAGACACTATATCTCATGCTCTAAGCTCTCAATGTT 690
Qy      203 G1YTYrAspValAlaSerArgLeuTYrProTIPGInCySpherIlePheHisAspValAsp 222
Db      691 GGTTTTCAGAACGCTTGAAGAGACTATGACTACACCGCTTTGCTTTAGTACGCGGAC 750
Qy      223 LeuLeuProGIuAspAspArgAenLeuTYrThrCysProIleGIInProArgHisMetSer 242
Db      751 CTCATTCATGATGATGACCATTAATGCGTACAGGTGTTTTCACAGCCAGCGCACTTCC 810
Qy      243 ValAlaIleAspIlySpherAenTYrIlySLeuProTYrSerAlaIlePheGIlyIleSer 262
Db      811 GTTGCAATGATGATGATTTGGATTTCAGCTTACTTATGTTTGAAGGTGCTCT 870
Qy      263 AlaLeuThrIlyAspHisLeuIlySlysIleAsnGIlyPheSerAspAspPheTIPGIlyTIP 282
Db      871 GCTTCTAGTAAACAACAGTTTCTTAACCATTAATGATTTCTTAATATTTTGGGCTGG 930
Qy      283 G1YGIlyGIuAspAspAspLeuValaThrArgThrSerMetAlaGIlyLeuIlyValSerArg 302
Db      931 GAGAGAGAAAGATGATGATTTTAAACAGATTGATTGTTTGAAGCATGCTCATATCTCGC 990
Qy      303 TYrProThrInIleAlaArgIlyIlySLeuMetIleIleHisSerThrGIuAlaThrAspPro 322
Db      991 CCAATGCTGTGTGGAGGTGTGCGATATCGCATCTCAGACCTCAGAGAAAGAAAGAA 1050
Qy      323 ValAsnIlySAspArgTYrIlySlysIleMetGIlyInThrIlySArgArgTIPThrArgAspGIly 342
Db      1051 CCCAATCTCAGAGGTTTGAACCAATTCACACACAAAGAGACAAATGCTCTCGATGGT 1110
Qy      343 LeuSerAsnLeuIlySTYrIlySLeuValAsnLeuGIlyLeuIlySProLeuTYrThrArgAla 362
Db      1111 TTGAACCTACCTACCTACCTACGAGGTGCTGATGATGACAGATATACCATGTATACCAATC 1170
Qy      363 ValValAspLeuLeuGIlySAspArgArgIlySLeuIlyArgArg 377
Db      1171 ACAGTGACATC-----GGGACACGAGCTAGAGATCCGTGCA 1206

RESULT 10
AA003464 standard; DNA; 2615 BP.
AC      AA003464;
XX      AC
XX      25-MAR-2003 (revised)
DT      15-AUG-1990 (first entry)
XX      XX
DE      cDNA encoding N-acetylglucosamine (beta 1-4) galactosyl transferase.
XX      Saccharide; teratocarcinoma F9 cell strain;
KM      N-acetylglucosamine (beta 1-4) galactosyl transferase; ss.
XX      XX
OS      Homo sapiens.
FH      Key Location/Qualifiers
FT      CDS 73..1269
FT      /tag= a
XX      XX
PN      JF02027987-A.
XX      XX
PD      30-JAN-1990.
XX      XX
PF      15-JUL-1988; 88JP-00176346.
XX      XX
PR      15-JUL-1988; 88JP-00176346.
XX      XX
PA      (GOKK-) GOKKO HOJIN.
XX      XX

```

```

DR      WPI; 1990-072976/10.
DR      P-PSDB; AAR05674.
XX      XX
PT      N-acetyl-glucosamine (beta 1-4) galactosyl transferase coding c-DNA -
XX      PT      useful for providing the transferase and hence for saccharide synthesis.
XX      XX
PS      Disclosure; Fig 2; 9pp; Japanese.
XX      XX
CC      The cDNA is very useful for prodn. of N-acetylglucosamine (beta 1-4)
CC      galactosyl transferase for in vitro synthesis of saccharide chains. See
CC      also AA003463. (Updated on 25-MAR-2003 to correct PD field.)
XX      XX
SQ      Sequence 2615 BP; 572 A; 694 C; 645 G; 704 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      7,84e-66      Length:      2615
Score:          668.00      Matches:      137
Percent Similarity: 54.69%      Conservative: 73
Best Local Similarity: 35.68%      Mismatches: 128
Query Match:    32.89%      Indels:      46
DB:             2      Gaps:      5

US-10-661-430-1 (1-383) x AA003464 (1-2615)
Qy      14 LeuLeuValLeuCyAlaValLeuLeuValIleAlaMetIleTYr----- 29
Db      148 CTCGCGCGCTGTGCGCTCGACCTTCCGCTCACCCCTGGTTACTACCTGCGCGCGC 207
Qy      30 -----LysIleProSerLeuTYrGIuAsnLeuThrIleGIlySerThrLeuIle 46
Db      208 GACCTAGCCGCGCTGCCCAACG-----GTGGAGCTTCACACCGCTG 252
Qy      47 AlaAspValAspAlaMetGIuAlaValIleGIlyAsnThrAla----- 60
Db      253 CAGGCGCGGTGCAACAGTCCGCCGCATCGGCACTCTCGGGAGAGCTCGGACCGGA 312
Qy      61 -----SerThrSer 63
Db      313 GGGGCGCGCGCGCTCTCTAGCGCTCTCCACCGCGCGCGCGCGCGCGACTCC 372
Qy      64 AspAspLeuLeuAspThrTrpAsnSerThrPheSerProIleSerGIuValAlaenGIInThr 83
Db      373 AGCCAGTCTGTGATCTGTGCGCTGCGCTGACCTGACACTGATGATCTCGGTC----- 423
Qy      84 SerPheMetGIuAspIleArgProIleLeuPheProAspAsnGIInThrLeuIlyPheCys 103
Db      424 -----CCAGTGCACCAACACACCGCACTGTGCGCGCGCTGC 462
Qy      104 AsnGIInThrProProHisLeuValGIlyProIleArgValPheLeuAspGIuPro---Asp 122
Db      463 CTTGAGGAGTCCCGCTGTGTGGGCCCATGCTGATGATGATTAACATGCTGTGAGC 522
Qy      123 PheIlyThrLeuGIlySleIlyTYrProAspThrHisAlaGIlyGIlyHisGIlyMetProIly 142
Db      523 CTGAGCTCTGTGCAAGAGCAAGAACCAATGTGAAGATGGCGCGCGCTATGCGCCAGG 582
Qy      143 AspCysValAlaArgHisArgValAlaIleIleValIProTYrArgAspArgIuAlaHis 162
Db      583 GACTGCTCTCTCTCAACAGTGGCCATCATCTCAATCCGCAACCGGACGAGAC 642
Qy      163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuValAlaGIInGIInLeuAspTYrAla 182
Db      643 CTCAGACTAGGCTATATTTTGGACCCAGCTCTGAGCGCGCAGGAGCTGACTATGGC 702
Qy      183 ILePheIleValIgluInValIalaenGIInThrPheAsnArgGIlyIlySLeuMetAenVal 202
Db      703 ATCTATGTTATCAACGAGCGGAGACACTATATCTCATGCTCTCAATGTT 762
Qy      203 G1YTYrAspValAlaSerArgLeuTYrProTIPGInCySpherIlePheHisAspValAsp 222
Db      763 GCTTTCAAGAACGCTTGAAGAGACTATGACTACACCGCTTTGCTTTAGTACGCGGAC 822
Qy      223 LeuLeuProGIuAspAspArgAenLeuTYrThrCysProIleGIInProArgHisMetSer 242

```


screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA gene sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

CC Sequence 3953 BP; 832 A; 1018 C; 981 G; 1122 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	1,65e-65	3953	123	52	93	1
Percent Similarity:	667.50					
Best Local Similarity:	65.06%					
Query Match:	45.72%					
	32.87%					

DB: 11 Gaps: 1

US-10-661-430-1 (1-383) x ACN45213 (1-3953)

OY 99 ThrLeuGlnPheCysAsnGlnThrProPheHisLeuValGlyProIleArgVal---Phe 117
 DB 551 TCGCTGCCAGCTTGGCTTACGAGATCCCGCTGCTGTTGGCCCATGCTGATGACTTT 610
 OY 118 LeuAspGluProAspPheThrLeuGluValIleTyProAspThrHisAlaGlyGly 137
 DB 611 AATATTGCTGTGTGATGTGAGCTTTTGGCAAGAACCCAGAGATAAAGACGGCGCG 670
 OY 138 HisGlyMetProLysArgPheCysValAlaArgHisArgValAlaIleIleValProTyArg 157
 DB 671 CGTTACTCCCGACAGACTGTGTCTCTCCACAAAGTGGCCATCATCCCATTCCT 730
 OY 158 AspArgGluAlaHisLeuArgIleMetLeuHisAsnLeuHisSerLeuValAlaValGln 177
 DB 731 AACCGGACGAGCATTTCAAAATACGTGCTGATATTGTCATCCCATCTTACGCCCG 790
 OY 178 GlnLeuAspTyAlaIlePheHisLeuValGlnValAlaAsnGlnThrPheAsnArgGly 197
 DB 791 CAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 850
 OY 198 LysLeuMetAsnValGlyTyArgPheValAlaSerArgLeuTyProTrpGlnCysPheIle 217
 DB 851 AAGCTGCTCAATATGCTTTCAGAGAGCTTGAAGAGCTATGATTAACAACCTGTTGTG 910
 OY 218 PheHisAspValAspLeuProGluAspAspArgAsnLeuTyThrCysProIleGln 237
 DB 911 TTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 970
 OY 238 ProArgHisMetSerValAlaIleAspLysPheAsnTyLysLeuProTySerAlaIle 257
 DB 971 CCACGGACATATTTCTTGGATGATGATGATGATGATGATGATGATGATGATGATG 1030
 OY 258 PheGlyGlyIleSerAlaLeuThrLysAspHisLeuLysLysIleAsnGlyPheSerArg 277
 DB 1031 TTTGGAGGTCT 1090
 OY 278 AspPheTrpGlyTyTrpGlyGlyGluAspAspLeuValAlaThrArgHisSerMetAlaGly 297
 DB 1091 AATTATTGGGGTTGGGAGAGAGAGATGACGACATTTTAAACGATTAAGTTACATAAGGC 1150
 OY 298 LeuLysValSerArgTyProThrGlnIleAlaArgTyLysMetIleLysHisSerThr 317
 DB 1151 AATGTAATATCATGCTCAATGCTGTAAGTGAAGAGGTGCGATATGATCCGGCATTTCAAGA 1210
 OY 318 GluAlaThrAsnProValAsnLysCysArgTyLysIleMetGlyGlnThrLysArgArg 337
 DB 1211 GACAGAGAAAATGAGCCCATCTCTAGAGGTTTGAACGGATCGACATTAACAAAGAAACG 1270

OY 338 TrpThrArgAspGlyLeuSerAsnLeuLysTyLysLeuValAsnLeuGluLeuLysPro 357
 DB 1271 ATGCCCTTCAGTGGTTTAACTACCTTACCAAGGTGTGAGATGATGATGATGATGATG 1330

OY 358 LeuTyThrArgAlaValAlaAspLeu 366
 DB 1331 TTATATACCCAAATCATGACGTGACATC 1357

RESULT 14

ID AAA53309 standard; cDNA, 1158 BP.

AC AAA53309;

DT 25-SEP-2000 (first entry)

DE Beta 1,4 galactose transferase (hGT) gene.

KW Beta 1,4 galactose transferase; hGT; transgenic plant; glycoprotein; ss.

OS Homo sapiens.

PN WO20034490-A1.

PD 15-JUN-2000.

PF 08-DEC-1999; 99WO-JP006881.

PR 09-DEC-1999; 98JP-00350584.

PA (SEKI/) SEKI T.

PA (FUJI/) FUJIYAMA K.

PI Seki T, Fujiyama K, Yoshida T;

DR WPI; 2000-431306/37.

PT P-PSDB; AAB03647.

PT Method for manufacturing a glycoprotein having human-type sugar chain, useful to produce glycoproteins for administration to animals and humans.

PS Example 1; Page 70-73; 79pp; English.

CC This sequence represents a human beta 1,4 Galactose transferase (hGT) gene. The gene is used in a method for manufacturing a glycoprotein with a human-type sugar chain. The method comprises obtaining a transformed plant cell, by introducing to a plant cell the glycosyltransferase gene and the gene of an exogenous glycoprotein, and cultivating the obtained transformed plant cell. Also included in the invention is a plant cell CC with a sugar chain adding mechanism which can conduct a transfer reaction of a galactose residue to a non-reducing terminal acetylglucosamine residue, where the sugar chain adding mechanism adds a sugar chain CC containing a core sugar chain and an outer sugar chain, where the core CC sugar chain comprises a number of mannose and acetylglucosamine, and CC where the outer sugar chain contains a terminal sugar chain portion with CC a non-reducing terminal galactose. The method can be used to produce a CC glycoprotein with a human-type sugar chain. The glycoproteins can be CC administered to animals, including humans. As the glycoprotein has a CC mammalian sugar chain (e.g. human type sugar chain) it is not antigenic CC as the glycosylation is of human type

SQ Sequence 1158 BP; 251 A; 347 C; 296 G; 264 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	3.98e-66	1158	137	73	128	5
Percent:	666.00					
Best Local Similarity:	54.69%					
Query Match:	35.68%					
	32.79%					

US-10-661-430-1 (1-383) x AAA53309 (1-1158)

```

Qy 14 LeuLeuValLeuCysAlaValLeuLeuValHisAlaMetIleTyr----- 29
Db 37 CTGTCGCGCGTCTGCGCTGCACTTGAGGTCAACCTGTTACTACTGCGTGGCCGC 96
Qy 30 -----LysIleProSerLeuTyrGluAsnLeuThrIleGlySerSerThrLeuIle 46
Db 97 GACCTGAGCGCGCTGCGCCCAACTG-----GTGGAAGTCTCACACCGCTG 141
Qy 47 AlaAspValAlaAspAlaMetGluAlaValLeuGlyAsnThrAla----- 60
Db 142 CAGGGCGGCTCGAAGAGTGGCCGCGCATGGGCAATCTCCGAGAGACTCCGACCGGA 201
Qy 61 -----SerThrSer 63
Db 202 GGGGCGCGCGCGCGCTCTCTAGGCGCTCTCCGACGGCGCGCGGTGGCGACTCC 261
Qy 64 AspAspLeuLeuAspThrTyrAsnSerThrPheSerProIleSerGluValAsnGlnThr 83
Db 262 AGCCCAAGTCGTGATTCGTGGCCCTGGCCGCTAGCAACTTGAACCTGGCTC----- 312
Qy 84 SerPheMetGluAspIleArgProIleLeuPheProAspAsnGlnThrLeuGlnPheCys 103
Db 313 -----CAGTGTCCCAACACACCGCACTGTGCTGCGCGCTGC 351
Qy 104 AsnGlnThrProProHisLeuValGlyProIleArgValPheLeuAspGluPro---Asp 122
Db 352 CCGAGAGAGTCCCGCTACTAGTGGGCCCATGTGATTTGAGCTTAAACATGCTGTGGAC 411
Qy 123 PheLeuThrLeuGluTyrIleTyrProAspThrHisAlaGlyGlyHisGlyMetProLys 142
Db 412 CTGAGAGTCGTGGCAAGCAGAACCCCAATGTGAGATGGGCGCGCTGTGCCCCCAGG 471
Qy 143 AspCysValAlaAsnHisArgValAlaIleIleValProTyrArgAspArgGluAlaHis 162
Db 472 GACTGCTCTCTCTCTCAAGAGTGGCCCATCTTCATCTTCGCAACCGCAGAGACAC 531
Qy 163 LeuArgIleMetLeuHisAsnLeuHisSerLeuLeuValLysGlnGlnLeuAspTyrAla 182
Db 532 CTCAGACTACGGCTATATTATTTCACACCGCTCTGAGGCGCAGCGCTGACTATGGC 591
Qy 183 IlePheIleValGluGlnValAlaAsnGlnThrPheAsnArgGlyLysLeuMetAsnVal 202
Db 592 ATCTATGTATCAACGAGCGGAGACATATATTCATGCTGCTAAGCTCTCAAGTGT 651
Qy 203 GlyTyrAspValAlaSerArgLeuTyrProGlnCysPheIlePheHisAspValAsp 222
Db 652 GGCCTTAAGAAGCGCTGAGAGACTATGACTACCTGCTGTGTGTTAGTGAAGTGAAC 711
Qy 223 LeuLeuProGluAspAspArgAsnLeuTyrThrCysProIleGlnProArgHisMetSer 242
Db 712 CTCATTCCAAATGATGACCATTAATGCGTACAGGTGTTTTCACAGCCACGCGCATTTCC 771
Qy 243 ValAlaIleAspLysPheAsnTyrLysLeuProTyrSerAlaIlePheGlyIleSer 262
Db 772 GTTGCAATGATGATTAATGATTGATCAGGCTTACCTTATGTTTGGAGGTGCTCT 831
Qy 263 AlaLeuThrLysAspHisIleLeuLysIleLeuGlnLysPheSerAsnAspPheThrGlyTyr 282
Db 832 GCTCTAAGTAAACAACAGTTTCTAACCAATCAATGATTTCTTAATTAATTAATGGGCTGG 891
Qy 283 GlyGlyGluAspAspLeuAlaThrArgThrSerMetAlaGlyLeuLysValSerArg 302
Db 892 GGAGGAGAAGATGATGACATTTTAAACGATTAGTTTAAAGCAGATGTATATCTCGC 951
Qy 303 TyrProThrGlnIleAlaArgTyrLysMetIleLysHisSerThrGluAlaThrAsnPro 322
Db 952 CCAATCTGTGTGCGGAGGTGTGCGATGATCCGCACTCAAGACACAAAGAAAATGAA 1011
Qy 323 ValAsnLysCysArgTyrLysIleMetGlyGlnThrLysArgGlyTyrThrArgAspGly 342
Db 1012 CCCAATCTCTAGAGAGTGTGACCGAATGACACAAAGGAGACAAATGCTCTGTGAAGGT 1071
Qy 343 LeuSerAsnLeuLysTyrLysLeuValAsnLeuGluLeuLysProLeuTyrThrArgAla 362

```

```

Db 1072 TTGAAGTCACTCACTACAGGTGTGATGACAGAGATACCAATGTATACCAAAATC 1131
Qy 363 ValValAspLeu 366
Db 1132 ACAGTGGACATC 1143

RESULT 15
ABK90350
ID ABK90350 standard; DNA; 1158 BP.
XX
AC ABK90350;
XX
DT 05-NOV-2002 (first entry)
XX
DE cDNA encoding human beta 1-4 galactose transferase (hgt).
XX
KW Gene; ss; beta1-4 galactose transferase; hgt; glycoprotein;
KW secretory production; sugar chain; acetylglucosamine; food; transgenic;
KW human.
XX
OS Homo sapiens.
XX
SY Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1158
FT /tag= a
FT /product= "beta 1-4 galactose transferase"
FT /BC_number= "2.4.1.38"
FT mutation replace(405,A)
FT /tag= b
FT mutation replace(408,A)
FT /tag= c
FT conflict replace(528,G)
FT /tag= d
FT /citation= [OTHER]
FT conflict replace(562,T)
FT /tag= e
FT /citation= [OTHER]
FT conflict replace(621..631,TA)
FT /tag= f
FT /citation= [OTHER]
FT conflict replace(1047,G)
FT /tag= g
FT /citation= [OTHER]
FT /note= "OTHER= Biochem. Biophys. Res. Commun., 157, 657-663 (1988)"
XX
XX WO200257468-A2.
XX
XX 25-JUL-2002.
XX
XX 18-JAN-2002; 2002WO-JP000361.
XX
XX 19-JAN-2001; 2001JP-00012519.
XX
XX (FUJI/) FUJIYAMA K.
XX (SEKI/) SEKI T.
XX
XX Fujiyama K, Seki T, Yoshida T;
XX
XX WPI; 2002-590749/63.
XX P-PSDB; ABG31511.
XX
XX Producing a glycoprotein with a human sugar chain comprises introducing
XX to a plant cell a gene of an enzyme capable of performing a transfer
XX reaction of a galactose residue, and a gene of a heterologous
XX glycoprotein.
XX
XX Example; Page 77-80; 82pp; English.
XX
XX This invention relates to a novel method for secretory production of a
XX glycoprotein with a human-type sugar chain. The method comprises

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 21:00:21 ; Search time 1076 Seconds
(without alignments)
8853.460 Million cell updates/sec

Title: US-10-661-430-2

Perfect score: 1152

Sequence: 1 atggtcttcgcatcttgcc.....acttccaagctgttttag 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications_NA_Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152	100.0	1152	7	US-10-661-430-2
2	192.6	16.7	1839	10	US-11-097-143-18554
3	167.6	14.5	2002	3	US-09-925-301-230
4	159.6	13.9	1062	7	US-10-287-226-287
5	159.6	13.9	1401	7	US-10-287-226-285
6	159.6	13.9	3931	8	US-10-757-262-111
7	159.6	13.9	3931	8	US-10-768-158-5
8	157.2	13.6	1853	3	US-09-745-763-84
9	155.4	13.5	1864	3	US-09-989-722-235
10	155.4	13.5	1864	3	US-09-989-723-235
11	155.4	13.5	1864	3	US-09-989-729-235
12	155.4	13.5	1864	3	US-09-989-727-235
13	155.4	13.5	1864	3	US-09-989-731-235
14	155.4	13.5	1864	3	US-09-989-732-235
15	155.4	13.5	1864	3	US-09-991-073-235
16	155.4	13.5	1864	3	US-09-990-442-235
17	155.4	13.5	1864	3	US-09-991-163-235
18	155.4	13.5	1864	3	US-09-993-604-235
19	155.4	13.5	1864	3	US-09-990-456-235
20	155.4	13.5	1864	3	US-09-989-721-235
21	155.4	13.5	1864	3	US-09-992-598-235
22	155.4	13.5	1864	3	US-09-989-293A-235
23	155.4	13.5	1864	3	US-09-989-735-235

24	155.4	13.5	1964	3	US-09-990-444-235	Sequence 235, App
25	155.4	13.5	1964	3	US-09-991-181-235	Sequence 235, App
26	155.4	13.5	1964	3	US-09-989-730-235	Sequence 235, App
27	155.4	13.5	1964	3	US-09-990-436-235	Sequence 235, App
28	155.4	13.5	1964	3	US-09-993-687-235	Sequence 235, App
29	155.4	13.5	1964	3	US-09-989-734-235	Sequence 235, App
30	155.4	13.5	1964	3	US-09-997-653-235	Sequence 235, App
31	155.4	13.5	1964	3	US-09-989-724-235	Sequence 235, App
32	155.4	13.5	1964	3	US-09-989-728-235	Sequence 235, App
33	155.4	13.5	1964	3	US-09-990-441-235	Sequence 235, App
34	155.4	13.5	1964	3	US-09-993-667-235	Sequence 235, App
35	155.4	13.5	1964	3	US-09-997-428-235	Sequence 235, App
36	155.4	13.5	1964	3	US-09-997-666-235	Sequence 235, App
37	155.4	13.5	1964	3	US-09-990-438-235	Sequence 235, App
38	155.4	13.5	1964	3	US-09-990-562-235	Sequence 235, App
39	155.4	13.5	1964	3	US-09-990-711-235	Sequence 235, App
40	155.4	13.5	1964	3	US-09-989-726-235	Sequence 235, App
41	155.4	13.5	1964	3	US-09-998-156-235	Sequence 235, App
42	155.4	13.5	1964	3	US-09-990-437-235	Sequence 235, App
43	155.4	13.5	1964	3	US-09-991-157-235	Sequence 235, App
44	155.4	13.5	1964	3	US-09-997-514-235	Sequence 235, App
45	155.4	13.5	1964	3	US-09-997-573-235	Sequence 235, App

ALIGNMENTS

RESULT 1
US-10-661-430-2
; Sequence 2, Application US/10661430
; Publication No. US20040086995A1
; GENERAL INFORMATION:
; APPLICANT: Cummings, Richard D.
; TITLE OF INVENTION: Beta 1, 4-N-ACETYLGLACTOSAMINYLTRANSFERASIS,
; TITLE OF INVENTION: NUCLEIC ACIDS AND METHODS OF USE THEREOF
; FILE REFERENCE: 7148.001
; CURRENT APPLICATION NUMBER: US/10/661,430
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: 60/411,242
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-661-430-2

Query Match 100.0%; Score 1152; DB 7; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCTTTTCGTCATTTGGCAGTCGCCAGACTCAAGTCGTTGCTGTACTTTGCGCTT	60
DB	1	ATGGCTTTTCGTCATTTGGCAGTCGCCAGACTCAAGTCGTTGCTGTACTTTGCGCTT	60
QY	61	CTTCTATTAGTTCATGCAATGTTTATAGATTCCATGCTTTAAGAACTTACTATC	120
DB	61	CTTCTATTAGTTCATGCAATGTTTATAGATTCCATGCTTTAAGAACTTACTATC	120
QY	121	GGCTCTCGACCCCTATTATGCGACGTCGACGATGAGAGGAGTCTCGGAATACGGCT	180
DB	121	GGCTCTCGACCCCTATTATGCGACGTCGACGATGAGAGGAGTCTCGGAATACGGCT	180
QY	181	TCACCTTGATGATCTACTTGATACGTGAATTCACGTTTTCACCGATTTCTGAAGTT	240
DB	181	TCACCTTGATGATCTACTTGATACGTGAATTCACGTTTTCACCGATTTCTGAAGTT	240
QY	241	AATCAGACTAGTTTATGAGAGCATTTGTCATTCCTGTCGCGAACAACGACTCTT	300
DB	241	AATCAGACTAGTTTATGAGAGCATTTGTCATTCCTGTCGCGAACAACGACTCTT	300

```

QY 301 CAATCTGTAATGACACCTCCCACTGTCGACCCATCCGTAATTCCTCATGAG 360
DB 301 CAATTTCTGTAATGACACCTCCCACTGTCGACCCATCCGTAATTCCTCATGAG 360
QY 361 CCCGACTTCAAACTCTCGAATAATCTATCCGACACGCAAGCCGCTGACATGGAATG 420
DB 361 CCCGACTTCAAACTCTCGAATAATCTATCCGACACGCAAGCCGCTGACATGGAATG 420
QY 421 CCTAAGGATGTTGTCAGAGGATCGGTTGCTATTATTGTCGCTATAGATCGGAA 480
DB 421 CCTAAGGATGTTGTCAGAGGATCGGTTGCTATTATTGTCGCTATAGATCGGAA 480
QY 481 GCACATTTGAGATAATGCTCCCAATTTGCACTGTTGCTGCGCAACAATTTGAC 540
DB 481 GCACATTTGAGATAATGCTCCCAATTTGCACTGTTGCTGCGCAACAATTTGAC 540
QY 541 TATGCAATTTTCAATTTGTCGACAGATGCGCAATTCAGCTTTTATGCGGGAATTAATG 600
DB 541 TATGCAATTTTCAATTTGTCGACAGATGCGCAATTCAGCTTTTATGCGGGAATTAATG 600
QY 601 AACGTTGATAGACAGATGATGACGCTTACCCATGCGAGCTTCACTTTCATGAT 660
DB 601 AACGTTGATAGACAGATGATGACGCTTACCCATGCGAGCTTCACTTTCATGAT 660
QY 661 GTGATTTTACGCGCGAAGATGACCGTACCTGTAACGCTGCAATTCACCAACGTCAT 720
DB 661 GTGATTTTACGCGCGAAGATGACCGTACCTGTAACGCTGCAATTCACCAACGTCAT 720
QY 721 ATGAGTGTAGCGATTCGATAATTTCAATTTAACTTTCATATTGCGGATCTTCCGCGA 780
DB 721 ATGAGTGTAGCGATTCGATAATTTCAATTTAACTTTCATATTGCGGATCTTCCGCGA 780
QY 781 ATGAGTGTAGCGATTCGATAATTTCAATTTAACTTTCATATTGCGGATCTTCCGCGA 840
DB 781 ATGAGTGTAGCGATTCGATAATTTCAATTTAACTTTCATATTGCGGATCTTCCGCGA 840
QY 841 GGTGGGCGGAGAGAGAGAGAGATTTGGCGACAGAACTGATGCTGACGTAAGAT 900
DB 841 GGTGGGCGGAGAGAGAGAGATTTGGCGACAGAACTGATGCTGACGTAAGAT 900
QY 901 TCAAGATATCCGACACAAATTTGACGATTAATAATGATTAAGCACTGACGAGGACG 960
DB 901 TCAAGATATCCGACACAAATTTGACGATTAATAATGATTAAGCACTGACGAGGACG 960
QY 961 AATCCAGTTAATAATGCGCTACAAATAATGAGGCGCAAGAGCGCGATGAGACAT 1020
DB 961 AATCCAGTTAATAATGCGCTACAAATAATGAGGCGCAAGAGCGCGATGAGACAT 1020
QY 1021 GACGCGCTAAGCAATCTGAATAGCTGTAATCTGTAATTTGAAGCTCTTCACT 1080
DB 1021 GACGCGCTAAGCAATCTGAATAGCTGTAATCTGTAATTTGAAGCTCTTCACT 1080
QY 1081 CGAGCGCTGCTGATTTGCTGCAAAAAGATGCGCGCGGAGCTGCGAAGGACCTTTCCA 1140
DB 1081 CGAGCGCTGCTGATTTGCTGCAAAAAGATGCGCGCGGAGCTGCGAAGGACCTTTCCA 1140
QY 1141 ACGTGTTTTAG 1152
DB 1141 ACGTGTTTTAG 1152

```

```

RESULT 2
US-11-097-143-18554
; Sequence 18554, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143

```

```

; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 4308
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18554
; LENGTH: 1839
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-18554
Query Match 16.7%; Score 192.6; DB 10; Length 1839;
Best Local Similarity 58.7%; Pred. No. 3,6e-51;
Matches 333; Conservative 0; Mismatches 234; Indels 0; Gaps 0;
QY 399 GCACGCGGTGACATGATGATGCTTAAGATTTGTTGCAAGGACGCTGCTATTAT 458
DB 600 GCGCCCTGTTGGGCGCTTGAAGCTGAAAACGCAATGCCAGCATCAGTGGCTATTGT 659
QY 459 TGTGCGCTAATGAGATGTAAGACATTTGAAATATGCTCCAAATTTGCACTGTT 518
DB 660 TGTGCGCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 719
QY 519 GCTGCGCAACAAACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 578
DB 720 TGTGATGAGAGAGCGGATGCGCTATGCAATTTTCAATTTGATGAGACACCAAGGAGCC 779
QY 579 GTTTAATCGCGGAACTAATGAACTTTGATGATGATGATGATGATGATGATGATGAT 638
DB 780 CTTTAATCGGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
QY 639 GCAGGCTTCACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 698
DB 840 GATTTGTTTATTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 899
QY 699 GTGTCAATTTCAACAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 758
DB 900 CTTGTCACGTCAGCGCGACACATGTCAGTGTATATGACACGCTGAATTCAGTTGCC 959
QY 759 ATATTGGGATTTTGGGCGAATTCAGTCACTTAACAAAGATCAGTGAAGAAATCA 818
DB 960 TTATGATCAATATTTTGAAGTGTTCCTGCAAGAGCGCTGACCACTTTCAGGCGCTAAA 1019
QY 819 TGAATTTGCAATGATTTTGGGTTGGGCGGAGAGAGAGATTTGGCGAGAAC 878
DB 1020 TGGGTTCTCAACTGCTTCTTGGCTGGGCGGCGAGATGACGATGCTCAACAGTT 1079
QY 879 ATGATGCTGACGTAAGTTTCAAGATATCCGACACAAATTTGACGATATTAATGAT 938
DB 1080 GAGGACGCAAACTTATCATATCAAGATATCCGCTCAACATATGCTCGTACAAATGCT 1139
QY 939 TAAGCATGACGAAAGCGAGATCC 965
DB 1140 GAAGCATGAGAAAGGCAATCC 1166

```

```

RESULT 3
US-09-925-301-230
; Sequence 230, Application US/09925301

```



```

? Patent NO. US20020052308A1
? GENERAL INFORMATION:
? APPLICANT: Rosen et al.
? TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
? FILE REFERENCE: PA106
? CURRENT APPLICATION NUMBER: US/09/925,301
? CURRENT FILING DATE: 2001-08-10
? PRIOR APPLICATION NUMBER: PCT/US00/05882
? PRIOR FILING DATE: 2000-03-08
? PRIOR APPLICATION NUMBER: 60/124,270
? PRIOR FILING DATE: 1999-03-12
? NUMBER OF SEQ ID NOS: 1694
? SOFTWARE: patentIn Ver. 2.0
? SEQ ID NO 230
? LENGTH: 2002
? TYPE: DNA
? ORGANISM: Homo sapiens
? OS-09-925-301-230

```

Query Match	14.5%;	Score 167.6;	DB 3;	Length 2002;
Best Local Similarity	54.4%;	Pred. No. 5.5e-43;		
Matches 360;	Conservative	0;	Mismatches 299;	Indels 3;
				Gaps 1;

QY	421	CCCTAAGGATGTTGTTGCAAGGACATCGGTGCTCTTATTTATGTGGCCCTTATGAGATTCGAA	480
Db	611	CTTGCAAGTTGTGAGCCCCCTGCTCCGAAACAGCAATCATTTGCTCTCATCGTCCCGGAG	670
QY	481	GCACTTTGAGAAATMACTGCTCACAAATTTTGCACTGTTGCTCGCCAAACAAATTTGGAC	540
Db	671	CACCACTGCGCCGTGCTGCTCTACCACTTGACCCCTTCTTGACGCGCAGACACTTGCT	730
QY	541	TATGCAATTTTTCATTTGTGAGCAAGTGGCCGCAATCAGACGTTTAAATCGGGGAAACTATG	600
Db	731	TATGGCACTTATGTGCATTCATCCAGGCTGGAAATGCAATTTTAAACGGGCAAACTGTTG	790
QY	601	AACGTTGGATACGACGTAGCATCAAGCCTCTACCCATGGCAGTGCTTCATCTTTCATGAT	660
Db	791	AACGTTGGGGTGAGAGAGGCCCTCGATGATGAGAGTGAGAGCTGCTGTTCTTGACAGAT	850
QY	661	GTGCATTTACTGCCCCGAAGATGACACCTGTAACCTGTACAGTGT---CCAATTCACACAGT	717
Db	851	GTGACACTCTTGGCCAGAAATGACCAACATCTGTATGTGTGTGACCCCGGGGACCCCGC	910
QY	718	CATATGAGTGTAGCGATCGATTAATTTCAATTATATACTTCATATTGCGGCATCTTGGC	777
Db	911	CATTTGCGCTTGTGCTATGAAACAAGTTTGGATATACGCTCCCGTACCCCGACTACTTGGGA	970
QY	778	GGAATCAGTGCATTACAAAAGATCACTGGAAGAAATCAATGATTTTGCATGATTTT	837
Db	971	GGATCTTCAGACATTACTCTCGAACAGTACCTGGAATGTGAATGTGCTTCCCAATGAATAC	1030
QY	838	TGGGGTTGGGGCGGAGAGGACGACGATTTTGGCGACGAAACATGATGGCTGACGCTGAAA	897
Db	1031	TGGGGCTGGGGTGTGAGAGATGACGACATTGCTTACAGAGTGGCGCTGTGGCTGGATGAG	1090
QY	898	GTTCCAAGATATCCGACACAAATTTGACGATATATAATGATTAAAGCTTCGACGGAAACG	957
Db	1091	ATCTCTCGGCCCCCACAATCTGTGTAGACACATATTAAGATGTGAAGCACCAGAGAGATPAG	1150
QY	958	ACGAATCCAGTTATTAATGTCCGCTACAAAATTAATGGCCCAAAGAACGCCGCAATGACA	1017
Db	1151	GGCAATAGAGAAATATCCCAAGATTTGACCTCTCGTGCCGTACCCAGAAATTCCTGGACG	1210
QY	1018	CGTAGCGGCTTAAGCAATCTGAAAGTATPAGCTCGTAAATCTGAAATGAAAGCCTCTAC	1077
Db	1211	CAAGATGGAGTAAGACTCACTGACATACACAGTTGTGCTGAGCTGAGAGCTGGGCTCTTTAT	1270
QY	1078	AC 1079	
Db	1271	AC 1272	

RESULT 4

```

US-10-287-226-287
Sequence 287, Application US/10287226
Publication No. US20040086875A1
GENERAL INFORMATION:
APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Berghs, Constance,
APPLICANT: Boldog, Ference,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eisen, Andrew,
APPLICANT: Ellerman, Karen,
APPLICANT: Gargolli, Esna A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weizhen,
APPLICANT: Kekuda, Rameeh,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Malvankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Ort, Tatiana,
APPLICANT: Padigara, Muralidhara,
APPLICANT: Paturajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark E.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderna, Steven K.,
APPLICANT: Spyrek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vernet, Corine A.M.,
APPLICANT: Zernusen, Bryan D.,
APPLICANT: Zhong, Mei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
CURRENT FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
PRIOR FILING DATE: 2001-11-20
Remaining Prior Application data removed - See File Wrapper or PALM
NUMBER OF SEQ ID NOS: 673
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 287
LENGTH: 1062
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

```

NAME/KEY: CDS
LOCATION: (2)..(1030)
US-10-287-226-287

Query Match 13.9%; Score 159.6; DB 7; Length 1062;
Best Local Similarity 54.3%; Pred. No. 1.7e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACACGCGCGGTGACATGGAATGCTTAAGATTGTGTGCAAGCATCTG 448
DB 297 ATTTAGATATTGACCGAGGGGTCTATTGAGGCCAAGACTTAACCCAGATGGAAG 356
QY 449 TTGCTATTATTGCGCTTATGAGATGCTGAGACCAATTGAGAAATATGCTCCACAT 508
DB 357 TGGCAGTCTTCATCTCTTCCGTAATGCGCATGAACTTCCAAATTTTCTTACATC 416
QY 509 TGACCTGTTGCTGCGCAACCAATTTGACATTTGCAATTTTCTGAGCAAGTGG 568
DB 417 TGATTCGAATGCTCCAGAACACGCGGTGGAATTTGGTTTATGTCATTGAAACAGACTG 476
QY 569 CGAATGACAGCTTTAATCGCGGGAACCTAATGAACTGATGCAAGTATGCAAGCC 628
DB 477 GCACACACCTTTTAAACCGTGCATGCTTTCAATGTGGCTTCAAGAGCCATGAAAG 536
QY 629 TCTACCATGACAGTGTCTTCTCATGATGTCGATTTTACTGCGCGAAGATGACGTA 688
DB 537 ACAGTGTCTGGACGTGTATCTTCCACGATGTGATCACTAAGTGAATGACCGGA 596
QY 689 ACCTGTACAGCTGTCCATTCACCAAGTCAATGATGATGAGATGATGATTTCAAT 748
DB 597 ACTATTAACGATGTGGAAGAAATGCAAGTCAATTTTGTGCAAGCTGATTAATACATGT 656
QY 749 ATAACTTCATATTGCGCGATCTTGGCGGAATCAAGTGCATTAACAAAGATCACTGA 808
DB 657 ATATTCTTCATATAAGAAATTTTGTGTGTGTAAGTGGCTGACAGTGAACAATTTA 716
QY 809 AGAAATCAATGATTTTGAATGATTTTGGGGTGGGGCGGAGAGACGACGATTTGG 868
DB 717 GAAAGATCAATGTTTCTTAATGCCCTTCTGGGGATGGGAGAGAAATGATGACCTTT 776
QY 869 CGACGACATGATGATGCTGACATGAAATTTCAATATCCGACCAATTTGACGAT 928
DB 777 GGAACAGAGTCACTATGCTGATATATATGTAACCAACAGAGGAGACTTTGGAAT 836
QY 929 ATAAATGATTAAGCACTGACGAGCAAGCAATTCAGTTAATTAATGCGCTACAAA 988
DB 837 ACAAGTCAATTTCTCTCA-TCAACATAGAGTGAAGTCCAGTT--TTAGGACGATTAAT 993
QY 989 TAAATGGCCAAACGAGCGCGATGACACGATGACGCGCTTAAGCAATCTGAAGTAAAG 1048
DB 894 TACTAAGGATATCCAAAGAGCGTCAATGATGATGACATGAACTTAATATATAGGC 953
QY 1049 TCGTAATCTGG 1060
DB 954 CAAATAATCTGG 965

RESULT 5
US-10-287-226-285
Sequence 285, Application US/10287226
Publication No. US20040086875A1

GENERAL INFORMATION:

APPLICANT: Agee, Michele L.,
APPLICANT: Alsobrook, John P.,
APPLICANT: Berghs, Constance,
APPLICANT: Boldog, Ferenc,
APPLICANT: Burgess, Catherine E.,
APPLICANT: Chant, John S.,
APPLICANT: Chaudhuri, Amitabha,
APPLICANT: DiPippo, Vincent A.,
APPLICANT: Edinger, Shlomit R.,
APPLICANT: Eilsen, Andrew,
APPLICANT: Ellerman, Karen,

APPLICANT: Gangolli, Esha A.,
APPLICANT: Gorman, Linda,
APPLICANT: Gerlach, Valerie,
APPLICANT: Ji, Weichen,
APPLICANT: Kekuda, Ramesh,
APPLICANT: Khramtsov, Nikolai,
APPLICANT: Li, Li,
APPLICANT: Malyankar, Uriel M.,
APPLICANT: MacDougall, John R.,
APPLICANT: Mezes, Peter S.,
APPLICANT: Miller, Charles E.,
APPLICANT: Millet, Isabelle,
APPLICANT: Ooi, Chean Eng,
APPLICANT: Padigaru, Muralidhara,
APPLICANT: Patnirajan, Meera,
APPLICANT: Rastelli, Luca,
APPLICANT: Rieger, Daniel K.,
APPLICANT: Rothenberg, Mark B.,
APPLICANT: Shenoy, Suresh G.,
APPLICANT: Spaderina, Steven K.,
APPLICANT: Spytek, Kimberley A.,
APPLICANT: Taupier, Jr., Raymond J.,
APPLICANT: Vernier, Corine A.M.,
APPLICANT: Zethusen, Bryan D.,
APPLICANT: Zhong, Mei,
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-480C
CURRENT APPLICATION NUMBER: US/10/287,226
PRIOR FILING DATE: 2002-11-04
PRIOR APPLICATION NUMBER: 60/334,421
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,392
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/360,148
PRIOR FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: 60/364,000
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/404,821
PRIOR FILING DATE: 2002-08-20
PRIOR APPLICATION NUMBER: 60/334,526
PRIOR FILING DATE: 2001-11-30
PRIOR APPLICATION NUMBER: 60/354,409
PRIOR FILING DATE: 2002-02-04
PRIOR APPLICATION NUMBER: 60/364,227
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: 60/334,027
PRIOR FILING DATE: 2001-11-28
PRIOR APPLICATION NUMBER: 60/331,641
Remainder Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 673
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 285
LENGTH: 1201
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (24)..(1169)
US-10-287-226-285

Query Match 13.9%; Score 159.6; DB 7; Length 1201;
Best Local Similarity 54.3%; Pred. No. 1.7e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACACGCGCGGTGACATGGAATGCTTAAGATTGTGTGCAAGCATCTG 448
DB 436 ATTTAGATATTGACCGAGGGGTCTATTGAGGCCAAGACTTAACCCAGATGGAAG 495
QY 449 TTGCTATTATTGCGCTTATGAGATGCTGAGACCAATTGAGAAATATGCTCCACAT 508
DB 496 TGGCAGTCTTCATCTCTTCCGTAATGCGCATGAACTTCCAAATTTTCTTACATC 555


```
RESULT 7
US-10-768-158-5
; Sequence 5, Application US/10768158
; Publication No. US20040204359A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Siles-Santiago, Immaculada
; APPLICANT: Karichet, Venkateswari
; APPLICANT: Riasof, Scott D
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 16386, 15402, 21165, 1423,
; TITLE OF INVENTION: 636, 12303, 21425, 27410, 38554, 38555, 55063, 57145, 59914,
; TITLE OF INVENTION: 94921, 16852, 33260, 58573, 30911, 85913, 14303, 16816,
; TITLE OF INVENTION: 17827 OR 32620
; FILE REFERENCE: MPIO3-012P1RNMNMIM
; CURRENT APPLICATION NUMBER: US/10/768,158
; CURRENT FILING DATE: 2004-01-30
; PRIOR APPLICATION NUMBER: US 60/444,781
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/452,291
; PRIOR FILING DATE: 2003-03-05
; PRIOR APPLICATION NUMBER: US 60/454,540
; PRIOR FILING DATE: 2003-03-13
; PRIOR APPLICATION NUMBER: US 60/478,805
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/491,048
; PRIOR FILING DATE: 2003-07-30
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 3931
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (298) ... (1446)
; US-10-768-158-5

Query Match      13.9%; Score 159.6; DB 8; Length 3931;
Best Local Similarity 54.3%; Pred. No. 3.4e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;
```

```
Db      1130 GAAAGATCAATGTTTCTTAATGCTTCTGCGGATGCGGAGAGAAATGATGACTTT 1189
Qy      869 CGACGAGATCATGATGCTGGAAGTTTCAAGATATCCGACACAAATTCACGAT 928
Db      1190 GGAACGAGTTTACATGCTGATGATATATGTAACGACGAGGAGCTTATGAGAAAT 1249
Qy      929 ATAAATGATTAAGCATCTGACGGAAGGACGATTCAGTTAATATCCGCTACAAA 988
Db      1250 ACAAGTCAATTCCTCA-TCACCATAGAGGTGAAGTTCAGTT--TTAGGACGGTATAAAT 1306
Qy      989 TAATGGCCCAACGAAAGCCGATGACACGTCAGCGCTTAAGCATCTGAAGTAAAC 1048
Db      1307 TACTAAGTATTTCCAGGAGCGTCGATCATGATGACTGAACAAATTTAATATATAGGC 1366
Qy      1049 TCGTAATCTGG 1060
Db      1367 CAAAAATACTGG 1378

RESULT 8
US-09-745-763-84
; Sequence 84, Application US/09745763
; Patent No. US20020065394A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; McCoy, John M.
; LaValle, Edward R.
; Collins-Racie, Lisa A.
; Evans, Cheryl
; Werberg, David
; Treacy, Maurice
; Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; ENCODING THEM
; NUMBER OF SEQUENCES: 219
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/745,763
; FILING DATE: 18-Jun-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 84:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1853 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 84:
US-09-745-763-84

Query Match      13.6%; Score 157.2; DB 3; Length 1853;
Best Local Similarity 54.4%; Pred. No. 1.3e-39;
Matches 360; Conservative 0; Mismatches 298; Indels 4; Gaps 2;
```

Db 499 CCTGAGGTTGAGCCCGCTCCGAGACGCCATCTTGTGCTCATGTGCTCCGAG 558
 Qy 481 GCACATTGAGATTAATGCTCCGACATTTGCACTGCTGTGCTGCGCAACAACATTTGAC 540
 Db 559 CACCACTTGGCTGCTGCTTACCACTGACACCTCTTCTTGAGAGCGCAAGCTTGTCT 618
 Qy 541 TATGCAATTTTCATTTGAGAGCAAGTGGCGAATCAGACGTTTAAATGCGGAACTAATG 600
 Db 619 TATGCACTATATGCTACCAAGCTGGAAATGAAATTTTAAAGGGGAAACCTGTTG 678
 Qy 601 AACGTTGATACAGCTAGCATAGCCCTCTTACCCATGCGAGTCTTCACTTTCAAT 660
 Db 679 AACGTTGGGAGTGGAGAGCGCTGCTGATGAGAGTGGAGCTGCTGTTCTTGACGAT 738
 Qy 661 GTGCAATTTACGCGCGAAGTACCGTAACTGTAACGGT---CCAAATTAACACAGT 717
 Db 739 GTGAGCTCTTGGCAGAAATGACCAATCTGATGTGTGTGATACCCCGGGGAGCCCGC 798
 Qy 718 CATATGAGTGTAGCGATCGATTAATTTCAATTAACCTTCAATTTGCGGATCTTGGC 777
 Db 799 CATGTTGCCGTTGCTATGAAACAAGTTGATACAGCTCCCGTACCCGAGTACTTGGGA 858
 Qy 778 GGAATGATGCTACTAACAAGATCACTGAGAGAAATCAATGATTTTCAATGATTTT 837
 Db 859 AAGATCTCAGCACTTCTCTGACAGTACCTGAGATGAATGGCTTCCCAATGAATAC 918
 Qy 838 TGGGGTTGGGGGCGAGAGACGATTTGGCGACAGAACTGATGCTGAGCTGAAG 897
 Db 919 TGGGGTGGGGTGTAGATGACGACTT-GCTACACAGGTGGCTGCTGGAGTGAAG 977
 Qy 898 GTTTCAGATATTCGACAAATTTGACATTAATTAATTAATTAAGACTCGACGAGAGCG 957
 Db 978 ATCTCTGCGGCCCCCACCACATCTGTAGACACTAATATGTGAAGACCGAGAGATTAAG 1037
 Qy 958 ACGAATCCAGTAAATATGCGCTTCAAAAATATGCGGCCAAGAGCGCGGAGTGA 1017
 Db 1038 GGCATAGAGAGAAATCCCAAGATTTGACCTCTGCTGCTGACCAAGATTTCTGAGAG 1097
 Qy 1018 CGTAGAGGCTTAAGCATCTGAAGTAACTGTAATCTGGAATTTGAAGCTCTCTAC 1077
 Db 1098 CAAAGTGGAGATGACATGACATGACATGCTGCTGCTGAGAGCTGCGGCTCTTAT 1157
 Qy 1078 AC 1079
 Db 1158 AC 1159

; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; FILE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2730PIC63
 ; CURRENT APPLICATION NUMBER: US/09/989,722
 ; PRIOR FILING DATE: 2001-11-19
 ; PRIOR APPLICATION NUMBER: 60/049787
 ; PRIOR FILING DATE: 1997-06-16
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/065186
 ; PRIOR FILING DATE: 1997-11-12
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066770
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/075945
 ; PRIOR FILING DATE: 1998-02-25
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/083322
 ; PRIOR FILING DATE: 1998-04-28
 ; PRIOR APPLICATION NUMBER: 60/084600
 ; PRIOR FILING DATE: 1998-05-07
 ; PRIOR APPLICATION NUMBER: 60/087106
 ; PRIOR FILING DATE: 1998-05-28
 ; PRIOR APPLICATION NUMBER: 60/087607
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087609
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087759
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 60/087827
 ; PRIOR FILING DATE: 1998-06-03
 ; PRIOR APPLICATION NUMBER: 60/088021
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088025
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088026
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088028
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088029
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088030
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088033
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088326
 ; PRIOR FILING DATE: 1998-06-04
 ; PRIOR APPLICATION NUMBER: 60/088167
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088202
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088212
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088217
 ; PRIOR FILING DATE: 1998-06-05
 ; PRIOR APPLICATION NUMBER: 60/088655
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: 60/088734
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088738
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088742
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088810
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088824
 ; PRIOR FILING DATE: 1998-06-10
 ; PRIOR APPLICATION NUMBER: 60/088826

PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25

PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 5,3e-39;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTTGTTCCAAAGCATGCTGTGCTATATATGTGCTTATAGATCTGAA 480
DB 461 CCTCAGAAATGTAAACCTTTACAGAGGTGCGCATCTGTTCCACCGAACAAGAG 520
QY 481 GCACATTGAGATATATGCTCCACATTTGCACTGCTGCGCAACACATTTGAC 540
DB 521 AAACACCTGATGTACTGCTGGAACATCTGCATCCCTTCTGAGAGGACGCTGAT 580
QY 541 TATGCAATTTTCAATGTGAGCAAGTGGCAATCAGACGTTTAAATGCGGAACTAATG 600
DB 581 TATGCAATCTAGCTATCAACAGCGCTGAAGTAAAGTTTATGAGCAAACTTGG 640
QY 601 AACGTTGATACAGAGTAGCATACGCTTACCCATGCGAGTCTTCACTTTATGAT 660
DB 641 AATGTGGCTATCTAAGAGCCCTCAAGAAAGAAATTTGGACCTTTATATTCACGAT 700
QY 661 GTGATTTACTGCGCGAAGATGACCGTAACCTGTACACGTCCTCAATTCACACGTCAT 720
DB 701 GTGACCTGTATCCAGAAATGTACTTTAACTTTAAGTGTAGAGACATCCCAAGCAT 760
QY 721 ATGAGTGTAGCGATCATTAATTAATTCAATTATTAATTTGGCGATCTTTGGCGGA 780
DB 761 CTGTTGTTGGAGGAACAGACACTGGGTACAGTTACGTTATGATATTTGGGGGT 820
QY 781 ATGATGTCATTAACAAAGATCCTGAGAAATCAATGATTTTTCGATGATTTTGG 840
DB 821 GTTACTGCCCCTAAGCAGAGAGAGTTTTCMAGTGATGATTTCTTAACAATCTGCG 880
QY 841 GGTGTGGGGGGAAGGACGAGATTTGGCAGAGAAATCATGATGCTGAGTGAAGTT 900
DB 881 GATTTGGGAGGCGAAGACATGATGACTCAGACTAGGGTTGAGCTCCAAAGATGAAATTT 940
QY 901 TCAAGATATCCGACACAATTTGACGATTAATTAATGATTAAGACCTTGACGAGCGACG 960
DB 941 TCCCGGCCCCCTGCTGAAGTGGGTAATATACATGCTTTCACACTAGACCAAGGC 1000
QY 961 AATCGATTATTAATGCGGCTCAAAATTAATGAGCCCAACGAGGCGCATGACGCT 1020

Db 1001 AATGAGTGAAGCAGACGATGAAGCTTTACACCAAGTGTACAGAGTCTGGAGACA 1060
Qy 1021 GACGGCTAGACATCTAGATCTGTAATCTGGAATTGAAGCTCTTACACT 1080
Db 1061 GATGGTTAGTGTCTTCTTATTAATTAATGTAATCTGTGAGACACATCTTATATATC 1120
Qy 1081 CGAGCCGTCGTGATTT 1097
Db 1121 AACATCAGAGTGATTT 1137

RESULT 10
US-09-989-723-235
; Sequence 235, Application US/09989723
; Patent No. US20020072092A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C62
; CURRENT APPLICATION NUMBER: US/09/989,723
; CURRENT FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03

; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088028
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088029
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088030
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088033
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088326
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088167
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088202
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088212
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088217
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/088655
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: 60/088734
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088738
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088742
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088810
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088858
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089440
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089599
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089600
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089801
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089907
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089908
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/089947
; PRIOR FILING DATE: 1998-06-19
; PRIOR APPLICATION NUMBER: 60/089948

PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 5.3e-39;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;
421 CTTAAGATTGTGTGCAAGCATCGTTCCTATTATGTGCGCCTATAGATCGTGA 480

Db 461 CCTCAGAAATGTAAGCTTTACAGAGGGTCGCCATCTCGTCCCAAGCAAGAG 520
Qy 481 GGACATTTAGAAATATATGCTCCAAATTTGCACTGTTGCTGCCCAACAAATTTGG 540
Db 521 AAACACCTATATGCTGCTGGAACATCTGCACTCTTCTCGAAGGACGCTGGAT 580
Qy 541 TATGCAATTTTCAATTTGAGGCAAGTGGCAATCAAGCTTTAATGCGGAAACTAATG 600
Db 581 TATGGCATTAAGTATCAACCAAGGCTGAAGGTAAATTAATTAATGACCAACTCTTG 640
Qy 601 AACGTTGATACGACCTAGCATCAAGCTCTCAATGCAATGCAATGCTTCAATTTAT 660
Db 641 AATGTCGCTATCTGAAGCCCTCAAGGAAGAAATTTGGAATGCTTTAATTTCCAGAT 700
Qy 661 GTGCATTTACTGCGCGAAGATGACCGTAACTGTACAGGTTCCAAATTAACACGCTAT 720
Db 701 GTGGACCTGTATCCGAGATGACTTTAACTTTACAGTGTAGAGAGATCCCAAGCAT 760
Qy 721 ATGAGGTAGCGATCGATTAATTTCAATTTAATTTCAATTTGCGCGATTTGCGCGA 780
Db 761 CTGCTGTTTGGCAGAAAGCACTGGGTACAGTTACGTTACAGTGATATTTGGGGAT 820
Qy 781 ATCAGTCACTAACAAAGATCACTGAAGAAATCAATGATTTTCAATGATTTTGG 840
Db 821 GTTACTGCCCTAAGCAGAGAGAGATTTTCAAGGTGAAGATTTCTTAACAACACTG 880
Qy 841 GGTGGGGGGGAGAGAGAGATTTGGGAGAGAAATCAATGCTGAGCTGAAGTT 900
Db 881 GGATGGGAGGCGAAGACATGCTCAAGCTGAGGTTGAGTCCAAAGATGAATTAAT 940
Qy 901 TCAAGATATCCGACACAAATTTGACGATTAATTAATTAATTAATTAATTAATTAAT 960
Db 941 TCCCGGCCCTGCTGAAGTGGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1000
Qy 961 AATCCAGTTAATTAATGCGCTAACAATTAATGAGGCAACGAGGCGGATGAGACAT 1020
Db 1001 AATGAGGTGAACGCAAGGATGAAGCTTTACAAAGTGTCAAGTGTGAGAAACA 1060
Qy 1021 GACGCGTAAAGCAATGAAATTAAGCTGTAATCTGTAATTTGAAGCTCTCACT 1080
Db 1061 GATGGTTGAGTATGTTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1120
Qy 1081 CGAGCCGCTGCTGATTT 1097
Db 1121 AACATCAAGTGATTT 1137

RESULT 11
US-09-989-279-235
Sequence 235, Application US/09989279
Patent No. US20020072496A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gunney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PLC56
CURRENT APPLICATION NUMBER: US/09/989,279
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678

PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090693
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 5.3e-19;

Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

421 CCTAAGATTGTGTGTCAGAGCATCGTGTCTATTATGTGCTTATAGATCGTGA 480
461 CCTCAGGAATGTAAGCTTACAGAGGCTGCCATCTCTGCCCGAACAAGAG 520
481 GCACATTTGAAATATGCTCCACATTTGCACTGCTCGCAACAACATTTGAC 540
521 AACACCTGATGTACCTGCTGGAACATCTGATCCCTCTGCAAGGCAAGCTGAT 580
541 TATGCAATTTCTATGTGAGCAAGTGGCAATCAGAGTTTATCGGGAAATCAT 600
581 TATGCACTTACCTGATCCACAGGCTGAAGTAAAGTTTATGAGCAAACTTTG 640
601 AACGTTGATACGACGATGATCAAGCTCTACCCATGCGAGTCTTCATCTTGCAT 660
641 AATGTGGCTATCTAGAAAGCCCTCAAGAAAGAAATTTGGAGCTCTTATATTCACGAT 700
661 GTGATTTATGCCCCGAAAGATGACCGTACCTGTACAGCTGTGCAATTCACAGCTCAT 720
701 GTGACCTGTATCCCGAATGATTTTAACTTTACAGTGTGAGAGCATCCCAAGCAT 760
721 ATGAGTTAGCATCGATGATAATTCATTTAACTTCATATTTGGGAGCTTGGGCGGA 780
761 CTGTGTGTTGAGAGAAAGACAGACTGGGTACAGTTACGTTACATGATATTTGGGGGT 820
781 ATCAGTGCATTAACAAAAGATCACCTGAGAAATCAATGATTTTGAATTTTGG 840
821 GTTACTGCTTAAAGCAGAGCAGATTTTTCAGAGTGAATGATTTCTTAACTACTGG 880
841 GGTGGGGCGGAGAGACAGCATTTTGGCGAGCAAGAACTGATGCTGAGCTGAAAGTT 900
881 GGAATGGGAGGCGGAAGAGATGACCTCAGACTCGGGTTGAGCTCAAAAGATTAATTT 940
901 TCAAGATATCCGACACAATTTGACGATTAATTAATGATTAAGCACTGAGAGGAGCAG 960
941 TCCCGGCCCCCTGAGTGGGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1000
961 AATCAGTTAATAATGCGCTACAAAATATATGGGCCAAAGAGCGCGATGAGACGAT 1020

1001 AATGAGTGAACGCAAGACGATGAGCTCTTACACCAAGTGTCAAGTCTGGAGACA 1060
1021 GAGCGCTTAAGAAATCTGAAGTAACTCGTAATTCGGAATTAAGCTCTTACACT 1080
1061 GATGGTGTGATGATGTTGTTCTTAAATTAATTAATTAATTAATTAATTAATTAATTA 1120
1081 CGAGCGGTGTCGATTT 1097
1121 AACATCAGTGTGATTT 1137

RESULT 12

US-09-989-727-235

Sequence 235, Application US/09989727

Patent No. US2002072497A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi J.

APPLICANT: Baker, Kevin P.

APPLICANT: Botstein, David

APPLICANT: Desnoyers, Luc

APPLICANT: Eaton, Dan L.

APPLICANT: Ferrara, Napoleone

APPLICANT: Fong, Sherman

APPLICANT: Gerber, Hanspeter

APPLICANT: Gerltzen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Absehn L.

APPLICANT: Kijavlin, Ivar J.

APPLICANT: Napier, Mary A.

APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann

APPLICANT: Scewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K.

APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

FILE REFERENCE: P2730P1C65

CURRENT APPLICATION NUMBER: US/09/989,727

CURRENT FILING DATE: 2001-11-19

PRIOR APPLICATION NUMBER: 60/049787

PRIOR FILING DATE: 1997-06-16

PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/065186

PRIOR FILING DATE: 1997-11-12

PRIOR APPLICATION NUMBER: 60/065311

PRIOR FILING DATE: 1997-11-13

PRIOR APPLICATION NUMBER: 60/066770

PRIOR FILING DATE: 1997-11-24

PRIOR APPLICATION NUMBER: 60/075945

PRIOR FILING DATE: 1998-02-25

PRIOR APPLICATION NUMBER: 60/078910

PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/083322

PRIOR FILING DATE: 1998-04-28

PRIOR APPLICATION NUMBER: 60/084600

PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/087106

PRIOR FILING DATE: 1998-05-28

PRIOR APPLICATION NUMBER: 60/087607

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087609

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087759

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 60/087827

PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19

PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 5,3e-39;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTGTCGAAGCATCGTGTGCTATTATGTCCTTAAGATCGTGA 480
DB 461 CCTCAGGAATGTAAAGCTTACAGAGGTCGCATCTCGTTCCTCCACCGGAACAGAG 520
QY 481 GCAATTTGAATTAATGCTCCCAATTTGCACTCGTGTGCCCAACAATTTGAC 540
DB 521 AAACACCTAATGACCTGCGAACAATCGCATCCCTTCTGAGAGGCGAGCGTGGAT 580
QY 541 TATGCAATTTTCACTGTGAGCAAGTGGCAATCAGAGCTTTAATGCGGGAATATG 600
DB 581 TATGCACTTACGCTACCTCCAGCGCTGAAGTAATAAGTTAATGAGCAAACTCTTG 640
QY 601 AACGTTGATACGACGTAGCATCAGCGCTTACCAATGCGAGCTTCACTTTCATGAT 660
DB 641 AATGTGGGCTATCTAGAAGCCCTCAAGAAAGAAATTTGGAGCTGTTATATTCAGAT 700
QY 661 GTGCAATTTACTGCCCAAGTACCGTAACCTGTACACGTTGCAATTCACCAAGTCA 720
DB 701 GTGGACCTGTATCCCGAATGACTTTTACCTTTACAAAGTGAAGAGCAATCCAAAGCAT 760
QY 721 ATGAGTGTAGCGATCGATTAATTCATTTAATCTTCATATTCGGGATCTTGGCGGA 780
DB 761 CTGCTGTGTGCGAGAACGACACTGGGTACAGTTACGTTACAGTGAATTTTGGGGGT 820
QY 781 ATCAGTGCATTAACAAAGATCACTGAAGAAATCAATGATTTTGAATGATTTTGG 840
DB 821 GTTACTGCTCTTACAGAGAGAGGATTTTTCAGAGTGAATGATTTCTTCACTACTGG 880
QY 841 GGTGGGGGCGGAAGACGATTTGGCGACGAAATCATGCTGATGCTGAACTGAAAGTT 900
DB 881 GGATGGGGAAGCGAAGACGATGACCTCAGACTGAGGTTGAGCTCCAAAGAAATTT 940
QY 901 TCAAGATATCCGACACAAATTTGACAGTATTAATGATTAAGCACTTCAGAGGAGCG 960
DB 941 TCCCGGCGCTGCTGCTGAATGGGTAATATACATGCTCTTCCACCTAAGACAAAGCG 1000
QY 961 AATCCAGTTAATAATGCGCTCAAAAATAATGGGCCAAAGCAAGCGCCGATGACAGT 1020
DB 1001 AATGAGGTGAACGACGAAGAGAGCTCTTACACCAAGTGCACGAGCTGGAAGACA 1060
QY 1021 GAGGCGCTAAGCAATCGAAGTATAGCTGTAAATTTGAAATTTGAAGCTCTCTACT 1080
DB 1061 GATGGGTGATGATGTTCTTTAATAATGATATGTTGGAACAACAATCTTTATATATC 1120
QY 1081 CGAGCGCTGCTGCAATTT 1097
DB 1121 AACATCACAGTGAATTT 1137

RESULT 13

US-09-989-731-235
Sequence 235, Application US/09989731
Patent No. US20020103125A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavini, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C70
CURRENT APPLICATION NUMBER: US/09/989,731
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824

?	PRIOR FILING DATE: 1998-06-10	?	PRIOR APPLICATION NUMBER: 60/090678
?	PRIOR FILING DATE: 1998-06-10	?	PRIOR FILING DATE: 1998-06-25
?	PRIOR FILING DATE: 1998-06-10	?	PRIOR APPLICATION NUMBER: 60/090690
?	PRIOR FILING DATE: 1998-06-11	?	PRIOR FILING DATE: 1998-06-25
?	PRIOR FILING DATE: 1998-06-11	?	PRIOR APPLICATION NUMBER: 60/090694
?	PRIOR FILING DATE: 1998-06-11	?	PRIOR FILING DATE: 1998-06-25
?	PRIOR FILING DATE: 1998-06-11	?	PRIOR APPLICATION NUMBER: 60/090695
?	PRIOR FILING DATE: 1998-06-11	?	PRIOR FILING DATE: 1998-06-25
?	PRIOR FILING DATE: 1998-06-11	?	PRIOR APPLICATION NUMBER: 60/090696
?	PRIOR FILING DATE: 1998-06-12	?	PRIOR FILING DATE: 1998-06-25
?	PRIOR FILING DATE: 1998-06-12	?	PRIOR APPLICATION NUMBER: 60/090862
?	PRIOR FILING DATE: 1998-06-16	?	PRIOR FILING DATE: 1998-06-26
?	PRIOR FILING DATE: 1998-06-16	?	PRIOR APPLICATION NUMBER: 60/090863
?	PRIOR FILING DATE: 1998-06-16	?	PRIOR FILING DATE: 1998-06-26
?	PRIOR FILING DATE: 1998-06-16	?	PRIOR APPLICATION NUMBER: 60/091360
?	PRIOR FILING DATE: 1998-06-16	?	PRIOR FILING DATE: 1998-07-01
?	PRIOR FILING DATE: 1998-06-16	?	PRIOR APPLICATION NUMBER: 60/091478
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR FILING DATE: 1998-07-02
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR APPLICATION NUMBER: 60/091544
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR FILING DATE: 1998-07-01
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR APPLICATION NUMBER: 60/091519
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR FILING DATE: 1998-07-02
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR APPLICATION NUMBER: 60/091626
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR FILING DATE: 1998-07-02
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR APPLICATION NUMBER: 60/091633
?	PRIOR FILING DATE: 1998-06-17	?	PRIOR FILING DATE: 1998-07-02
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR APPLICATION NUMBER: 60/091978
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR FILING DATE: 1998-07-07
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR APPLICATION NUMBER: 60/091982
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR FILING DATE: 1998-07-07
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR APPLICATION NUMBER: 60/092182
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR FILING DATE: 1998-07-09
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR APPLICATION NUMBER: 60/089908
?	PRIOR FILING DATE: 1998-06-18	?	PRIOR FILING DATE: 1998-06-18
?	PRIOR FILING DATE: 1998-06-19	?	PRIOR APPLICATION NUMBER: 60/089947
?	PRIOR FILING DATE: 1998-06-19	?	PRIOR FILING DATE: 1998-06-19
?	PRIOR FILING DATE: 1998-06-19	?	PRIOR APPLICATION NUMBER: 60/089948
?	PRIOR FILING DATE: 1998-06-19	?	PRIOR FILING DATE: 1998-06-19
?	PRIOR FILING DATE: 1998-06-19	?	PRIOR APPLICATION NUMBER: 60/089952
?	PRIOR FILING DATE: 1998-06-19	?	PRIOR FILING DATE: 1998-06-19
?	PRIOR FILING DATE: 1998-06-22	?	PRIOR APPLICATION NUMBER: 60/090246
?	PRIOR FILING DATE: 1998-06-22	?	PRIOR FILING DATE: 1998-06-22
?	PRIOR FILING DATE: 1998-06-22	?	PRIOR APPLICATION NUMBER: 60/090252
?	PRIOR FILING DATE: 1998-06-22	?	PRIOR FILING DATE: 1998-06-22
?	PRIOR FILING DATE: 1998-06-22	?	PRIOR APPLICATION NUMBER: 60/090254
?	PRIOR FILING DATE: 1998-06-22	?	PRIOR FILING DATE: 1998-06-22
?	PRIOR FILING DATE: 1998-06-23	?	PRIOR APPLICATION NUMBER: 60/090349
?	PRIOR FILING DATE: 1998-06-23	?	PRIOR FILING DATE: 1998-06-23
?	PRIOR FILING DATE: 1998-06-23	?	PRIOR APPLICATION NUMBER: 60/090355
?	PRIOR FILING DATE: 1998-06-23	?	PRIOR FILING DATE: 1998-06-23
?	PRIOR FILING DATE: 1998-06-23	?	PRIOR APPLICATION NUMBER: 60/090429
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090431
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090435
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090444
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090445
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090472
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090535
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090540
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090542
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR APPLICATION NUMBER: 60/090557
?	PRIOR FILING DATE: 1998-06-24	?	PRIOR FILING DATE: 1998-06-24
?	PRIOR FILING DATE: 1998-06-25	?	PRIOR APPLICATION NUMBER: 60/090676
?	PRIOR FILING DATE: 1998-06-25	?	PRIOR FILING DATE: 1998-06-25

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
 Best Local Similarity 51.8%; Pred. No. 5.3e-39;
 Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

Qy	421 CCTAAGATTGTTGGCAAGCATGTTGCTATTATTTGSCCTATAGAGATGCTGAA	480
Db	461 CCTCAGAAATGTAAGCTTTACAGAGGTCGCACTCTCTCCCAACGGAAGAG	520
Qy	481 GCACATTTGAGAAATATGCTCCAAATTTGCACTGCTGCTGCCAACAACATTTGAC	540
Db	521 AAACACCTGATGTACTGCTGTGACATCTGATCCCTTCTCAGAGGACGAGCTGAT	580
Qy	541 TATGCAATTTTATTGTTGAGCAAGTGGCAATCGACGTTTAATCGCGGAACTAATG	600
Db	581 TATGCAATCTACGTATCCACAGGCTGAAGGTAAAGTTTAATCGAACCAACTCTTG	640
Qy	601 AACGTTGATTAAGAGATGATCAAGCCCTTACCATGAGAGTGTCTTTCATGAT	660
Db	641 AATGCGGCTATCTTGAAAGCCCTCAAGAGAAATTTGGACTGCTTTATTTCCAGAT	700
Qy	661 GTGCAATTTACTGCCGGAAGATGACGTAACCTGTACAGGTGTCAATTCAACCAAGTCAT	720
Db	701 GTGCACTGTTACCGGAATGACTTTAACTTTCAAGTGTGAAGAGATCCCAAGAT	760
Qy	721 ATGAGTATGAGCATGATTAATTTCAATTAACTTTCAATTTCGGCATCTTTCGGCGA	780
Db	761 CTGAGTGTGGGAGAGAAACGCACTGGGTACAGTTACGTAAAGATATTTGGGGGT	820
Qy	781 ATCAGTGACCTAAACAAAGATCACCTGAAGAAATCAATGATTTTCGATGATTTTGG	840
Db	821 GTTACTGCCCTTAAGAGAGAGAGATTTTCAAGGTGATGATTTCTTAACTAACTACTGG	880
Qy	841 GATTGGGGGAGAGAGAGAGATTTGGGAGAGAAATCATGATGATGATGATGATGAT	900
Db	881 GGATGGGGGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT	940
Qy	901 TGAAGTATCGACCAATTTGACGATTAATAATGATTAAGATCTGACGAGAGAGAG	960
Db	941 TCCGCGCCCTGCTGAAAGTGGTAAATTAATTAATGATGATGATGATGATGATGAT	1000

PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089948
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/089952
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090246
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090252
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090254
 PRIOR FILING DATE: 1998-06-22
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090355
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090431
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090435
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090444
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090472
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090535
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090540
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090542
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090557
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090676
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090678
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090690
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090694
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090695
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090696
 PRIOR FILING DATE: 1998-06-25
 PRIOR APPLICATION NUMBER: 60/090862
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091478
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091544
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091626
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091633
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091978
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07
 PRIOR APPLICATION NUMBER: 60/092182
 PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
 Best Local Similarity 51.8%; Pred. No. 5.3e-39;
 Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTTGTTGCAAGGAGCATGTTGCTATTAATTGTCCTTAAGAGTCGTGA 480
 DB 461 CCTCAGAGATTTAAAGCTTTAAGAGGTCGTCCTTCCCAACGAGAGAG 520
 QY 481 GCACATTTGAGATTAATGCTCAATTTGCACTGTCCTGCGCAACCAATTTGAC 540
 DB 521 AAACACCTGATGTATCTGCTGGAACATCTGATCCCTTCGACAGGAGAGCTGAT 580
 QY 541 TATGCAATTTTATTTGAGAGCAAGTGGCAATCAGCTTTAATTCGCGGAACTAAG 600
 DB 581 TATGGATCTTACGTATCCACAGGCTGAAGGTAAAGTTAATGACCAAACTCTTG 640
 QY 601 AACGTTGATACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 641 AATGTCGCTATCTGAAAGCCCTCAAGAGAAATTTGGAATGCTTTATTTCCAGAT 700
 QY 661 GTGATTTTACGTCGCGGAGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 701 GTGACCTGCTATCTGAAAGCCCTCAAGAGAAATTTGGAATGCTTTATTTCCAGAT 760
 QY 721 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 DB 761 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 820
 QY 781 ATCAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 821 GTTATCTGCTTACGAGAGAGATGATGATGATGATGATGATGATGATGAT 880
 QY 841 GGTGAGGAG 900
 DB 881 GATGAGGAG 940
 QY 901 TCAATATCTGACCAATTTGATGATGATGATGATGATGATGATGATGAT 960
 DB 941 TCCGAGGAG 1000
 QY 961 AATCAGTTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1020
 DB 1001 AATGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1060
 QY 1021 GACGAGCTTAAGATCTGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1061 GATGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGAT 1120
 QY 1081 CGAGGCTGCTGATTT 1097
 DB 1121 AACATCAGATGATTT 1137

RESULT 15
 US-09-991-073-235
 Sequence 235, Application US/09991073
 Patent No. US20020127576A1
 GENERAL INFORMATION:
 APPLICANT: Ashkenazi, Avi J.
 APPLICANT: Baker, Kevin P.
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Baton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Fong, Sherman
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, J. Christopher
 APPLICANT: Gurney, Austin L.
 APPLICANT: Kijavir, Ivar J.
 APPLICANT: Napier, Mary A.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Matanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C15
CURRENT APPLICATION NUMBER: US/09/1991.073
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10

PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676

PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 5.3e-39;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

421 CCTAAGGATTGTGTCAGAGCATGCTGTGCTATTATTTGTCCTATAGAGTGTGAA 480
461 CCTCAGGAATGTAAAGCTTTACAGAGGTGCGCCATCTGTTCCCAACGGAACAGAG 520
481 GCACATTTGAGATATATGCTCCACATTTGCACTGTTGCTGCCAACAACATTTGAC 540
521 AAACACCTGATGACCTGTGAGACATCTGCATCCCTTCTGACAGAGGACGCTGAT 580
541 TATGCAATTTTCAATTTGAGAGCAAGTGGCAATCAGACGTTTATGCGGAAACTAATG 600
581 TATGGCATCTAGCTATCCACAGGCTGAAGGTAAAGTTTATGAGCCAAACTCTTG 640
601 AACGTTGATAGACGTACATCAGCCCTCTACCCATGGCAGTCTTCAATTCATGAT 660
641 AATGGGCTATCTTAAGACCCCTCAAGAGAAATTTGGACTCTTATATTCACAGAT 700
661 GTGCAATTTACTGCCCGAAGATGACCGTAACTGTACAGCTTCATTCACCAACGTCAT 720
701 GTGACCTGTGACCGAGATGATGATTTAACTTTACAGTGTAGAGAGATCCCAAGCAT 760
721 ATGAGTINAGCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
761 CTGGTGTGTGGCAGAAAGCACTGGGTACAGGTACGTACAGTATATTTTGGGGGT 820
781 ATCAGTGCACTAACAATAATCACTGAAAGAAATCAATGATTTTCAATGATTTTGG 840
821 GTTACTGCTTAAGCAGAGAGCTTTTCAAGGTAAATGATCTCTTAACAACATACTG 880
841 GGTGGGGCGAGAGAGACGATTTGGGACGAGAACTCATGCTGAGCTGAAGTT 900
881 GGATGGGAGGAGAGACGATGACCTCAGACTCAGGTTGAGCTCCAAGAGATGAATTT 940
901 TCAAGTATCCAGACAAATTTGACAGATTAATAATGATTAAGCATCTGACGAGAGCAG 960
941 TCCCGCCCTGCTGAGAGTGGGTAAATATATCAATGCTCTTCACTAGAGCAAAAGGC 1000

961 AATCCAGTTAATAATGCGCTACAAAATATATGCGCAACGAGCGCATGACACGT 1020
1001 AATGAGGTGAAGCAGAAACGATGAGAGCTTTACACCAAGTCTACAGTCTGGAGACA 1060
1021 GAGGCTTAAGCAATCTGAAGTATAGCTGTAAATCTGAAATTTGAAGCTCTTACACT 1080
1061 GATGGTTGAGTATGTTGTTCTTAATAATTAATGATCTGGAAACAAATCTTTATATATC 1120
1081 CGAGCGTCTGCTGATTT 1097
1121 AACATCAGATGATTT 1137

Search completed: January 9, 2006, 00:54:49
Job time : 1081 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 18:50:05 : Search time 243 Seconds
(without alignments)
8426.959 Million cell updates/sec

Title: US-10-661-430-2

Perfect score: 1152

Sequence: 1 atgccttcctcattcattgcgc.....acttccacagctgttttag 1152

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/1/COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5/COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A/COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B/COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H/COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS/COMB.seq:*
7: /cgn2_6/ptodata/1/ina/RE/COMB.seq:*
8: /cgn2_6/ptodata/1/ina/BACKFILES1.seq:*
9: /cgn2_6/ptodata/1/ina/BACKFILES1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	191	16.6	1189	3	US-09-270-767-11472
2	167.6	14.5	1938	3	US-09-673-395A-140
3	167.6	14.5	1970	3	US-09-949-016-2875
4	161.2	13.0	3830	3	US-09-949-016-679
5	159.6	13.9	3832	3	US-09-949-016-2359
6	155.4	13.5	1964	3	US-09-991-181-235
7	155.4	13.5	1964	3	US-09-990-444-235
8	155.4	13.5	1964	3	US-09-997-333-235
9	155.4	13.5	1964	3	US-09-992-598-235
10	155.4	13.5	2151	3	US-09-949-016-1767
11	155.4	13.5	2167	3	US-09-949-016-541
12	155.4	13.5	2280	2	US-09-055-097-4
13	155.4	13.5	2280	2	US-09-373-902-4
14	150	13.0	2265	3	US-09-949-016-4375
15	150	13.0	4646	3	US-09-949-016-680
16	139.6	12.1	1265	2	US-08-446-777-1
17	139.6	12.1	2304	2	US-08-446-777-5
18	139.6	12.1	2304	2	US-08-446-777-7
19	134	11.6	1023	3	US-09-118-464-34
20	134	11.6	1023	3	US-10-132-652A-34
21	134	11.6	1116	3	US-09-118-464-33
22	134	11.6	1116	3	US-10-132-652A-33
23	134	11.6	1119	3	US-09-949-016-2195
24	134	11.6	1119	3	US-09-949-016-2195

25	134	11.6	2027	3	US-09-118-464-1	Sequence 1, Appl1
26	134	11.6	2027	3	US-10-132-652A-1	Sequence 1, Appl1
27	131.6	11.4	494	3	US-09-270-767-27061	Sequence 27061, A
28	108.8	9.4	541	3	US-09-404-879A-108	Sequence 108, App
29	108.8	9.4	541	3	US-09-338-933-108	Sequence 108, App
30	108.8	9.4	541	3	US-09-215-681-108	Sequence 108, App
31	108.8	9.4	541	3	US-09-215-681-108	Sequence 108, App
32	108.8	9.4	541	3	US-09-215-681-108	Sequence 108, App
33	108.8	9.4	541	3	US-09-667-857-108	Sequence 108, App
34	108.8	9.4	541	3	US-10-198-053-108	Sequence 108, App
35	69.4	6.0	500	3	US-09-827-271-108	Sequence 108, App
36	69.4	6.0	500	3	US-09-270-767-1933	Sequence 1933, App
37	66.8	5.8	26313	3	US-09-270-767-17215	Sequence 17215, A
38	66.8	5.8	26313	3	US-09-949-016-16117	Sequence 16117, A
39	66.8	5.8	81384	3	US-09-949-016-12422	Sequence 12422, A
40	58	5.0	601	3	US-09-949-016-81352	Sequence 81352, A
41	58	5.0	65518	3	US-09-949-016-12421	Sequence 12421, A
42	58	5.0	65518	3	US-09-949-016-14101	Sequence 14101, A
43	57.2	5.0	601	3	US-09-949-016-31230	Sequence 31230, A
44	57.2	5.0	601	3	US-09-949-016-31231	Sequence 31231, A
45	57.2	5.0	601	3	US-09-949-016-81354	Sequence 81354, A

ALIGNMENTS

RESULT 1						
US-09-270-767-11472						
; Sequence 11472, Application US/09270767						
; Patent No. 6703491						
; GENERAL INFORMATION:						
; APPLICANT: Homburger et al.						
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster						
; FILE REFERENCE: File Reference: 7326-094						
; CURRENT APPLICATION NUMBER: US/09/270,767						
; NUMBER OF SEQ ID NOS: 1999-03-17						
; SOFTWARE: PatentIn Ver. 2.0						
; SEQ ID NO 11472						
; LENGTH: 1189						
; TYPE: DNA						
; ORGANISM: Drosophila melanogaster						
US-09-270-767-11472						
Query Match 16.6%; Score 191; DB 3; Length 1189;						
Best Local Similarity 58.6%; Pred. No. 1.2e-51;						
Matches 332; Conservative 0; Mismatches 235; Indels 0; Gaps 0;						
QY	399	GCACGCCGCTGACATGGAATGCTTAAGATGTGTGCAAGCAGTCGTTGCTATTAT	458			
DB	508	GGCCCTGTGTGCGCTTGCAGCTTGAATCTGCAATGCCAGATCCACGTGCTATTTGT	567			
QY	459	TGTGCTTATAGATGCTGGAAGCATTTGGAATTAATGCTCAATTTGCTGCTGTT	518			
DB	568	TGTGCTTGTGCGCATGATGAGCCATCTATTAATCTTTCTGCGCAATCAGCCAT	627			
QY	519	GCTGCCCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG	578			
DB	628	TCTGATGAGGAGCGCCATGCTATGCAATTTCAATTTGAGGAGCAACGAGGAGCC	687			
QY	579	GTTTATGCGCGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG	638			
DB	688	CTTCAATGCGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	747			
QY	639	GCAGTCTTCAATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG	698			
DB	748	GGATTTGTTTATTTATTCAGATGATGATGATGATGATGATGATGATGATGATGATG	807			
QY	699	GTTTCAATTCACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG	758			
DB	808	CTGTTCAGTGAAGCGGCAACATGCTGATGATGATGATGATGATGATGATGATGATG	867			
QY	759	ATATTCGCGATCTTGTGCGGATGATGATGATGATGATGATGATGATGATGATGATG	818			

Db 868 TTATCATCAATTTTGGAGGTGTTCCGCAATGACCGGTGAGACACTTTCAGCCCGTAA 927
Qy 819 TGGATTTTGAATGATTTTGGGGTGGGGCGGAGAGAGACGATTTGGCCAGGAAC 878
Db 928 TGGGTTCTCAAACTGTTCTTTGGCTGGGGCGGAGAGATGAGACATGCTCCAAAGGTT 987
Qy 879 ATGATGAGCTGAGCTGAAAGTTTCAAGATATCCGACCAAAATTGACGATATAAATGAT 938
Db 988 GAAGACGCCCACTTATTCATTCAGAGTATCCGGTCAACATAGCCCGCTCAAGATGCT 1047
Qy 939 TAAGCACTGACGAGGAGCGAGCAATCC 965
Db 1048 GAAGCATCAGAAAGAAAGGCCAATCC 1074

RESULT 2

US-09-673-395A-140
; Sequence 140, Application US/09673395A
; Patent No. 6620923
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
; FILE REFERENCE: ALBRE-12
; CURRENT APPLICATION NUMBER: US/09/673, 395A
; CURRENT FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 637
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 140
; LENGTH: 1938
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-673-395A-140

Query Match 14.5%; Score 167.6; DB 3; Length 1938;

Best Local Similarity 54.4%; Pred. No. 7.1e-44; Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

Qy 421 CCTAAGGATTGTGTGCAAGGACATCGTGTGCTATTATTTGTCCTATAGAGATCGTGA 480
Db 573 CTGCGAGGTGTGAGCCCCGCTCCGAAACGACATATTGTGCTCATCGCCGGAG 632
Qy 481 GCACATTGAGAAATATGCTTCCACAAATTTGCACTGTTGCTGCCAAACAAATGGAC 540
Db 633 CACCACCTGCGCTGCTCTTACCACTGACACCCCTTCTTGACGCCGACAGCTTGCT 692
Qy 541 TATGCAATTTTCAATTTGGAAGCAAGTGGCGAATGAGCGTTTATGCGGGAACATAAG 600
Db 693 TATGAGCATCTATGATCATCCACAGGCTGAAATGAAATTTTACAGGGCAAACTTTG 752
Qy 601 AACGTGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 753 AAGTTTGGGTGGAGAGGCGCTGATGAGAAAGTGGAGCTGCTGTTCTTGAGAT 812
Qy 661 GTGATTTTACTGCCCCGGAAGATGACCTGTAACGTGT---CCAATTCAACAGCT 717
Db 813 GTGACCTCTTGGCAGAAATGACCAATCTGTATGTGTGACCCCGGAGACCCGCG 872
Qy 718 CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
Db 873 CATGTGCTGTTGTATGAAACAGTTTGGATACAGCTCCCGTACCCCGACGTACTTGA 932
Qy 778 GAATCAGTGCATCAAAAGATCACTGAAAGAAATCAATGATTTTGAATGATTTT 837
Db 933 GGAATCTCAGACATTAATCTTCCGACCAATGATGATGATGATGATGATGATGAT 992
Qy 838 TGGGGTTGGGGCGGAGAGACGATTTGGCGACGAAACATGATGCTGAGCTGAAA 897

Db 993 TGGGGCTGGGGTGTGAGATGACGATTTGCTACCGAGGTGCCCTGGCTGGATGAAG 1052
Qy 898 GTTTCAGATATCCGACCAAAATTTGACAGATATAAATGATTAAGCATCGACGAGCG 957
Db 1053 ATCTTCGGCCCCCAGATCTGTAGACACTTATAGATGTGAAGCAGCAGAGATGAAG 1112
Qy 958 ACGAATCCAGTTAATTAATGACGCTACAAATATATGGCCAAACGAAAGCCGATGACA 1017
Db 1113 GCGAATGAGAAATATCCAGATTTGACCTGTCGTAACCAAGATTTCTTGACG 1172
Qy 1018 CGTACGCTTAAAGCAATGTGAATTAAGTCTGTAATCTGGAATTAAGCTCTTAC 1077
Db 1173 CAAGATGGATGAACTCACTGATCAATACAGTTGTGGCTGAGAGCTGGGGCTTTAT 1232
Qy 1078 AC 1079
Db 1233 AC 1234

RESULT 3

US-09-949-016-2875
; Sequence 2875, Application US/09949016
; Patent No. 6812338
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for windows Version 4.0
; SEQ ID NO 2875
; LENGTH: 1970
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2875

Query Match 14.5%; Score 167.6; DB 3; Length 1970;

Best Local Similarity 54.4%; Pred. No. 7.2e-44; Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

Qy 421 CCTAAGGATTGTGTGCAAGGACATCGTGTGCTATTATTTGTCCTATAGAGATCGTGA 480
Db 956 CTGCGAGGTGTGAGCCCCGCTCCGAAACGACATATTGTGCTCATCGTCCGGAG 1015
Qy 481 GCACATTGAGAAATATGCTTCCCAATTTGCACTGTTGCTGCCAAACAAATGGAC 540
Db 1016 CACCACCTGCGCTGCTCTTACCACTGACACCCCTTCTTGACGCCGACAGCTTGCT 1075
Qy 541 TATGCAATTTTCAATTTGGAAGCAAGTGGCGAATGAGCGTTTATGCGGGAACATAAG 600
Db 1076 TATGCAATCTATGATCATCCACAGGCTGAAATGAAATTTTACAGGGCAAACTTTG 1135
Qy 601 AACGTGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 1136 AAGTTTGGGTGGAGAGGCGCTGATGAGAAAGTGGAGCTGCTGTTCTTGACGAT 1195
Qy 661 GTGATTTTACTGCCCCGGAAGATGACCTGTAACGTGT---CCAATTCAACAGCT 717
Db 1196 GTGACCTCTTGGCAGAAATGACCAATCTGTATGTGTGACCCCGGAGACCCGCG 1255
Qy 718 CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777
Db 1256 CATGTGCTGTTGTATGAAACAGTTTGGATTAAGCTCCCGTACCCCGACGTACTTGA 1315
Qy 778 GAATCAGTGCATCAAAAGATCACTGAAAGAAATCAATGATTTTGAATGATTTT 837

Db 1316 GGAAGTCTCAGCACTACTCTGACCGATGACATGATGATGCTTCCCAATGAATAC 1375
Qy 838 TGGGGTTGGGGGAGAGAGACGATTTGGGACGAGAACTCGATGGCTGACCTGA 897
Db 1376 TGGGGCTGGGGGAGAGATGACGATTTGCTACCGAGGGTGGCTGGCTGGATGAAG 1435
Qy 898 GTTTCAGATATCCGACACAAATTTGACGATTAATAATGATTAAGCATCCGACGAGCG 957
Db 1436 ATCTCTGGCCCCCAGCATCTGTAGACACTATAGATGATGAAGCAGGAGAAATAG 1495
Qy 958 ACGAATCCAGTTAATTAATGCGCTACAAATTAATGAGCGCAAGAGCGCGATGACA 1017
Db 1496 GGCATGAGAGAAATCCCGACAGATTTGACCTCTGCTCGTACCGAATTCCTGAGCG 1555
Qy 1018 CGTACGCGCTTACGACATTTGAAGTTAAGCTCTGTAATCTGGAATGAGCTCTAC 1077
Db 1556 CAAGATGGAGTAACTCACTGACATACGATTTGCTGGCTGAGAGCTGGGGCTCTTAT 1615
Qy 1078 AC 1079
Db 1616 AC 1617

RESULT 4
US-09-949-016-679
; Sequence 679, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 679
; LENGTH: 3830
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-679

Query Match 14.0%; Score 161.2; DB 3; Length 3830;
Best Local Similarity 54.5%; Pred. No. 1.3e-41;
Matches 366; Conservative 0; Mismatches 303; Indels 3; Gaps 2;
Qy 389 ATCCGACACGACGCGGTGACATGGAATGCTTAAGATTTGTTGCAAGGATCTG 448
Db 652 ATTGATATTTAGAGCAGGGGCTCATTTGAGGCCAAAGACCTTAACCAATGGAAG 711
Qy 449 TTGCTATTATTTGCTTATGAGATCGTGAAGCACATTTGAGATTAATGCTCCAAAT 508
Db 712 TGGCAGTTCTCATTTCTTTCCGTAATCGCCATGAACATCTCCAAATTTTCTTACATC 771
Qy 509 TGCACCTGTTGCTCGCAACAAACAATTGACATGATGCAATTTTCAATGTGAGCAAGTGG 568
Db 772 TGATTTCAATGCTCCAGAGACAGCGGCTGAATTTGCGTTTATGTGTTGAACAGACTG 831
Qy 569 CGAATCAGAGCTTTATCGCGGAAACTAATGACGTTGATGACGATGATGACGCGC 628
Db 832 GCAACAACCTTTTAACTGAGATCTTTTCAATGTGAGCTTCAAGAGGCGCATGAAG 891
Qy 629 TCTACCATGAGCAGTCTTCAATCTTTTCAATGATGATGATGATGATGATGATGATG 688
Db 892 ACAGTGTCTGGAGCTGTGTATCTTCCAGATGTGATCATCTACCTGAATAATGACCGGA 951

Qy 689 ACCTGACAGCTGTCTCAATTCACCACTCATATGATGATGATGATGATGATGATGATG 748
Db 952 ACTATTAGAGATGTGAGAAATGCGACGTCAATTTTGGTCAAGCTGATTAATATCATGT 1011
Qy 749 ATAACTTCAATATTCGCGATCTTTCGCGGATGATGATGATGATGATGATGATGATG 808
Db 1012 ATATCTTCCATATTAATAATTTTGTGTGATGATGATGATGATGATGATGATGATG 1071
Qy 809 AGAAATCAATGATTTTCAATGATTTTGGGGGTTGGGGGAGAGAGCGAGATTTGG 868
Db 1072 GAAAGATCAATGATTTTCTTAATGCTTCTGAGAGAGAGAGAGAGATGATGATGATG 1131
Qy 869 CGACGAGAACATCGATGCTGCTGACATGAAGTTTCAAGATTCGACACAAATTTGACGAT 928
Db 1132 GGAACAGATTCACATCTGATGATTAATGTAACGACGAGGAGAGCTTGAAGAAAT 1191
Qy 929 ATAAATGATTAAGATCTGACGAGAGGAGGAGATTCGATTAATGATGATGATGATG 988
Db 1192 ACAGTCAATTCCTCA-TCACCATGAGATGATGATGATGATGATGATGATGATGAT 1248
Qy 989 TAATGGCGCAACGAGAGCGGATGACGACGATGACGCTGAAGATGATGATGATGATG 1048
Db 1249 TACTAAGTATTCAGAGAGGCTCATGATCATGATGATGATGATGATGATGATGATG 1308
Qy 1049 TCGTAATCTGG 1060
Db 1309 CAAAATACTGG 1320

RESULT 5
US-09-949-016-2359
; Sequence 2359, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2359
; LENGTH: 3832
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2359

Query Match 13.9%; Score 159.6; DB 3; Length 3832;
Best Local Similarity 54.3%; Pred. No. 4.2e-41;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;
Qy 389 ATCCGACACGACGCGGTGACATGGAATGCTTAAGATTTGTTGCAAGGATCTG 448
Db 652 ATTGATATTTAGAGCAGGGGCTCATTTGAGGCCAAAGACCTTAACCAATGGAAG 711
Qy 449 TTGCTATTATTTGCTTATGAGATCGTGAAGCACATTTGAGATTAATGCTCCCAAT 508
Db 712 TGGCAGTTCTCATTTCTTTCCGTAATCGCCATGAACATCTCCAAATTTTCTTACATC 771
Qy 509 TGCACCTGTTGCTCGCAACAAACAATTGACATGATGCAATTTTCAATGTGAGCAAGTGG 568
Db 772 TGATTTCAATGCTCCAGAGACAGCGGCTGAATTTGCGTTTATGTGTTGAACAGACTG 831
Qy 569 CGAATCAGAGCTTTATCGCGGAAACTAATGACGTTGATGATGATGATGATGATGATG 628
Db 832 GCAACAACCTTTTAACTGAGATCTTTTCAATGATGATGATGATGATGATGATGATG 891

PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090682
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090683
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360

PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 76-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

421 CCTAAGATTGTTGGCAAGCATGTTGCTATTATTGTCCTTAATGAGATGCTGAA 480
461 CCTCAGATGTAAGCTTTACAGAGGTCGCTCTGTTCCCAACGGAACAGAGAG 520
481 GGACATTTGAGATATAGTCCAAATTTGCACTGTTGCTGCTGCCAAACAATTTGAC 540
521 AAACACCTGATCTACCTGTAACATCTGCACTCCCTTCTCAGAGGAGCACTGGAT 580
541 TATGCAATTTGATTTGAGGCAAGTGGCAATGAGCACTTAATTCGCGGAACCTAATG 600
581 TATGATATAGTATCATCCACAGGCTGAAGTTAAAGTTAATGAGCCAACTCTTG 640
601 AAGTTGATACAGATGATCAAGCTCTTACCATGAGCACTGTTCTTCTTATGAT 660
641 AATGTTGGCTATCTGAAGCCCTCAAGGAAGAAATTTGGGACTGTTTATTTCCAGAT 700
661 GTGCAATTTACGCGCGGAAGATGACCGTAACTGTACAGTGTCCAAATTAACAGCTCAT 720
701 GTGACCTGTGACCGGAGATGATCTTAACTTTACAGTGTGAGAGATGATCCAAAGAT 760
721 ATGAGTTAGCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
761 CTGTGTGTTGAGAGAAACAGCACTGGGTACAGTTACGTTACAGTGATATTTGGGGGT 820
781 ATCAGTGCATTAACAAGATCACTGAAGAAATCAATGATGATTTTGCATGATTTTGG 840
821 GTTACTGCCCTAAGCAGAGAGATTTTCAAGTGAATGATTTCTTAACAACCTACTGG 880
841 GGTGGGGGGGAGAGAGACGATTTGGGAGAGAAATCGATGGCTGAGCTGAAGATT 900
881 GGATGGGAGGAGAGACGATGACCTGAGCTGAGGTTGATCCAAAGAAATTAATTT 940
901 TCAAGATATCCAGACAAATTCAGATTAATTAATTAATTAATTAATTAATTAATTAAT 960
941 TCCCGCCCTCTGTAAGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1000
961 AATCAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020
1001 AATGAGTGAACGCAAGACGATGAAGCTTTACACCAAGTGTCAAGACTGTGAGAAACA 1060
1021 GACGCGCTTAAGCAATCTGAAGTAACTGCTTAATTCGGAATTAAGCTCTCTCAACT 1080
1061 GATGGTGTGATGATGTTGTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1120
1081 CGAGCGCTGCTGATTT 1097
1121 AACATCAGAGTGATTT 1137

RESULT 7
US-09-990-444-235

Sequence 235, Application US/09990444
Patent No. 6930170
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730P1C19
CURRENT APPLICATION NUMBER: US/09/990,444
PRIOR FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033

PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23

PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;

Best Local Similarity 51.8%; Pred. No. 7e-40; Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTTGTCAGAGCATCGTGTATATATGTCCTATAGAGATCGTGA 480
DB 461 CCTCAGAGATGTAAAGCTTTACAGAGAGGTGCGCATCTGTTCCACGGAACAAGAG 520
QY 481 GCACATTGAGATATATGTCCTCAATTGTCATGCTGTGTCGCCAACAACATTTGAC 540
DB 521 AAACACCTTAGTACCTGCGAACAATCTGCATCCCTTCTGAGAGGACAGCTGGAT 580
QY 541 TATGCAATTTTCAATGTGAGCAAGTGGCAATCAGCGTTTATATGCGGGAATTAAG 600
DB 581 TATGCAATTTTCAATGTGAGCAAGTGGCAATCAGCGTTTATATGCGGGAATTAAG 640
QY 601 AACGTTGATACAGCTAGCATCAGCGCTTACCATGCGAGTCTTCAATTTTCATGAT 660

DB 641 AATGCGGCTATCTAGAACCCCTCAAGAGAAATTTGGAGCTTTATATTCACGAT 700
QY 661 GTGCAATTTCTGCCCCGGAANATGACCGTACCTGTACAGTGTCCATTTAACCACTCAT 720
DB 701 GTGCACTGTGACCCGAGATGACTTTAACCTTTACAGTGTGAGAGCATCCCAAGCAT 760
QY 721 ATGAGTGTAGCAGTATGATTAATTTCAATTTCCATTTCCGCGATCTTCGCGGA 780
DB 761 CTGCTGTGTGAGAGAGACGACCTGGGTACAGTTACGTTACAGTGTATTTTGGGGCT 820
QY 781 ATCAGTGTCTACCAAAAGATCCTGAGAAATTCATGTATTTTGAATGATTTTGG 840
DB 821 GTTACTGCCCTTAAAGAGAGAGAGCTTTTCAAGTGAATGATTTCTTAACTACTG 880
QY 841 GGTGGGGGAGAGAGAGAGATTTTGGGAGAGAAATCATGATGCTGTGATGAAGTT 900
DB 881 GGATGGGAGGCGAAGACATGACCTCAGCTCAGGTTGACCTCCAAAGATGAATTT 940
QY 901 TCAAGTATCCGACACAAATTTGACGATTAATTAATGATTAAGCACTGACGAGGAG 960
DB 941 TCCCGCCCCCTGCTGAGATGAGTAAATTAATTAATGCTTTCCACACTAGACCAAGGC 1000
QY 961 AATCAATTAATTAATGCGCTTACAAATATGAGCCAAAGAGCCGATGACACGT 1020
DB 1001 AATGAGTGAACGCAAGACGATGAAGCTTTACACCAAGTGTACAGTGTGAGAA 1060
QY 1021 GAGCGCTTAAAGCAATCTGAGTAACTGCTTAATTTGGAATTTGAAGCTCTTCACT 1080
DB 1061 GATGGTTGAGTATGTTGTTCTTATTAATTAATTAATGATCTGTGAAACAAATCTTTATATATC 1120
QY 1081 CGAGCGTGTGATTT 1097
DB 1121 AACATCAGATGATTT 1137

RESULT 8
US-09-997-333-235
Sequence 235, Application US/09997333
Patent No. 6953836
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Geritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Guiney, Austin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC27
CURRENT APPLICATION NUMBER: US/09/997,333
CURRENT FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17

PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 7e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

421 CCTAAGATTGTTGTCAGAGCATGCTGTTGCTATTATTGTCCTATTAGAGATGTTGAA 480
461 CCTCAGGAATGTAAGACTTTACAGAGGGTCCGATCCTGTTCCCAAGGAAAGAGAG 520
481 GGACATTTGAGAAATATGTCACAAATTTGCACTGCTGCTGCCAAACAAATTTGAC 540
521 AAACACCTGATGATGACCTGCTGGAACATCTGCAATCCCTCTCCAGAGGAGAGAGCTGAT 580
541 TATGCAATTTTCACTATTGAGAGCAAGTGGCAATCAGACGTTTAATGCGGAAACTATG 600
581 TATGGCATTAAGTATCACTCAACAGGCTGAAGTAAAAATTTATGAGCCAACTCTTG 640
601 AAGTTTGATACGAGCTAGCATCAAGCTCTCAACCAATGCAAGTCTTCAATCTTTCATGAT 660
641 AATGTTGGCTATCTAGAAAGCCCTCAAGAAAGAAATTTGGAGATGCTTTATTTCCAGAT 700
661 GTTCGATTTACTGCCCCGAAATGACCCCTAACCTGTACACTGTCCAAATTTAACCACTCAT 720
701 GTGAGCTGCTAACCCGAAATGACTTTTAACTTTAAGATGAGAGATCCCAAGCAT 760
721 ATGAGTATGCGATGATTAATTTCAATTATTAATTTCCATTTGCGCGATCTTTCGCGGGA 780
761 CTGGTGTGTTGCGAGAAAGCACTGGGTACAGTTACGTTACAGTGGATATTTTGGGGGT 820
781 ATCAGTGCACTAACAAAGATCACTGAAGAAATCAATGATTTTGAATGATTTTGG 840
821 GTTACTGCCCCTAAGCAGAGAGAGTCTTTCAAGGTGAATGATTTCTTAACAACACTCTGG 880
841 GGTGGGGGGGAGAGACGACGATTTGGGAGAGAAACATGATGGCTGAGCTGAAAGTT 900
881 GGATGGGGGGGGAACACATGACCTCAGACTCAGGGTTGAGCTCCAAAGATGAAATTT 940
901 TCAAGATATCCGACCAATTTGACGATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
941 TCCCGCCCCCTGCTGAATGGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1000
961 AATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
1001 AATGAGGTGAACGAGAAACGATGAAGCTCTTAACAAAGTGTACAGATCTGGAAGACA 1060
1021 GACGGCCTAAGCAATTTGAAGTATTAAGCTGTAAATTTGGAATTTGAAGCTCTTAACACT 1080
1061 GATGGGTGAAGTATTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1120
1081 CGAGCGCTGCTGATTT 1097
1121 AATCATCAGTGTGATTT 1137

RESULT 9

US-09-992-598-235
Sequence 235, Application US/09992598
Patent No. 6956108
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi J.
APPLICANT: Baker, Kevin P.
APPLICANT: Bolstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Auestin L.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2730PIC20
CURRENT APPLICATION NUMBER: US/09/992,598
CURRENT FILING DATE: 2001-11-14
PRIOR APPLICATION NUMBER: 60/049787
PRIOR FILING DATE: 1997-06-16
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/075945
PRIOR FILING DATE: 1998-02-25
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087607
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088021
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04

PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355

PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090662
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090663
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 7e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTTGGTGAAGCATCGTTGCTATTATTGTGCCCTTAATGAGATCGTGA 480
Db 461 CCTCAGGAATGTAAAGCTTTACAGAGGGTGCATCTCGTCCCAACGGAACAAGAG 520
QY 481 GCACATTGGAGTAATGCTCCACATTGGCATCTGCTGCCCAACAACATGGAC 540
Db 521 AAACACCTGATGTACCTGCTGGAACATCTGCATCCCTCTCGACAGGCGAGCTGGAT 580
QY 541 TATGCAATTTTCATTGTGGAGCAAGTGGCAATCAAGCTTTATATGCGGGAATATATG 600
Db 581 TATGCAATTTTCATTGTGGAGCAAGTGGCAATCAAGCTTTATATGCGGGAATATATG 640
QY 601 AACGTTGATACGACGTAAGCATCAGCGCTTACCATGGACGATGCTTCAATCTTATGAT 660

Db 641 AATGGGCTATCTGAAGCCCTCAAGAGAAATGGAATGCGCTTATATTCACGAT 700
Qy 661 GTCCATTTACTCCCGGAAGATGACCTGTAACCTGTACAGTGTCAATTAACCACTGAT 720
Db 701 GTGACCTGTGATCCGAGAAATGACTTTAACTTTCAAGTGTGAGAGAGATCCCAAGAT 760
Qy 721 ATGATGTAGCGATGATTAATTTCAATTTCAATTTCCATTTCCGCGCATCTTCGCGGA 780
Db 761 CTGTGTGTGGAGAGAGACGACTGGGTACAGTTACGTTCACAGTGTGATTTTGGGGGT 820
Qy 781 ATCACTGACCTAACAAAGATCACTGAAAGAAATCAATGATTTTGAATGATTTTGG 840
Db 821 GTTACTGCCCCCTAAGAGAGAGAGATTTTCAAGGTGAATGATTTCTTAACAACACTG 880
Qy 841 GGTGGGGCGAGAGAGACGATTTTGGGAGAGAAATCATGATGCTGATGCTGATGAAAGT 900
Db 881 GGATGGGGAGGAGAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAAATGAAAT 940
Qy 901 TCAAGATATCCGACAAATTCAGATATTAATGATTAATGACTCGACGAGAGGAGC 960
Db 941 TCCCGCCCCCTGCTGAGTGGTAAATATACATGTTCTTCAACTAGAGACAAAGC 1000
Qy 961 AATCCAGTTAATTAATGCCGCTACAAATTAATGGCCAAACGAGCGCGATGACAGT 1020
Db 1001 AATGAGGTGAACGCAAGAGATGAAGCTCTTAACACCAAGTGTCAAGTCTGGAAGAC 1060
Qy 1021 GACGCGCTAAGCAATCTGAAGTATAGCTGTAAATCTGGAATTTGAAGCTCTTCACT 1080
Db 1061 GATGGGTGAAGTATGTTCTTAAATTAATGATGCTGTGGAACACAAATCTTTATATATC 1120
Qy 1081 CGAGCGGTGCTGATTT 1097
Db 1121 AACATCACAGTGATTT 1137

RESULT 10
US-09-949-016-1767
; Sequence 1767, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1767
; LENGTH: 2151
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1767

Query Match 13.5%; Score 155.4; DB 3; Length 2151;
Best Local Similarity 51.8%; Pred. No. 7.3e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

Qy 421 CCTAAGATTGTGTGCAAGGATCGTGTCTATTATTTGCCCCCTAAGAGATCGTGA 480
Db 653 CCTCAGAAATGTAAGCTTTACAGAGGGTCCGATCTGTTCCCAAGCAAGAGAG 712
Qy 481 GACATTTAGATATGTCACAAATTTGCACTGTGTCTGCCCAAAACAAATTTGAC 540
Db 713 AAACACGTATGTACTCTGGAACATCTGCAATCCCTTCTGCAAGGAGAGAGCTGGAT 772

Qy 541 TATGCAATTTTCAATTTGTGAGCAAGTGGCAATCAGACGTTTAATTCGGGAAACTAATG 600
Db 773 TATGGCACTACGTATCACCAGGCTAAGGTAAAGTTTAATTCAGCCAAACTCTTG 832
Qy 601 AACGTTGATACGATGATGATCAAGCTCTTACCCAGTGGTTCATCTTTTATGAT 660
Db 833 AATGGGCTATCTTAAGAGCCCTCAAGAGAAATTTGGGACTGCTTTATATTTCCACGAT 892
Qy 661 GTCCATTTACTGCCCGGAAGATGACCGTAACTGTACAGTGTCCAAATCAACCACTGAT 720
Db 893 GTGACCTGTATCCGAGAAATGACTTTAACTTTAACAAGTGTGAGAGACATCCCAAGAT 952
Qy 721 ATGATGTAGCGATGATTAATTTCAATTTAACTTCCATTTTCCGCGATCTTCGCGGA 780
Db 953 CTGTGTGTGGAGAGAAACAGCACTGGGTACAGTTACGTTACAGTGTATTTTGGGGGT 1012
Qy 781 ATCACTGACCTAACAAAGATCACTGAAAGAAATCAATGATTTTGAATGATTTTGG 840
Db 1013 GTTACTGCCCCCTAAGAGAGAGACAGTTTTCAGGTGAATGATTTCTTAACAACACTG 1072
Qy 841 GGTGGGGCGAGAGAGAGATTTTGGCGAGAGAAATCATGATGCTGACTGAAAGT 900
Db 1073 GGATGGGAGGAGAGAGATGACCTCAGACTCAGGGTTGAGCTCAAAAGATGAAAT 1132
Qy 901 TCAAGATATCCGACAAATTCAGATATTAATTAATTAAGCACTCGACGGAACGAGC 960
Db 1133 TCCCGCCCCCTGCTGAGTGGTAAATATACATGTTCTTCAACTAGAGACAAAGC 1192
Qy 961 AATCCAGTTAATTAATGCCGCTACAAATTAATGGCCAAACGAGCGCGATGACAGT 1020
Db 1193 AATGAGGTGAACGCAAGAGATGAAGCTCTTAACACCAAGTGTCAAGTCTGGGAACA 1252
Qy 1021 GACGCGCTAAGCAATCTGAAGTATAGCTGTAAATCTGGAATTTGAAGCTCTTCACT 1080
Db 1253 GATGGGTGAAGTATGTTCTTAAATTAATGATGCTGTGGAACACAAATCTTTATATATC 1312
Qy 1081 CGAGCGGTGCTGATTT 1097
Db 1313 AACATCACAGTGATTT 1329

RESULT 11
US-09-949-016-541
; Sequence 541, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 541
; LENGTH: 2167
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-541

Query Match 13.5%; Score 155.4; DB 3; Length 2167;
Best Local Similarity 51.8%; Pred. No. 7.4e-40;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

Qy 421 CCTAAGATTGTGTGCAAGGATCGTGTCTATTATTTGCCCCCTAAGAGATCGTGA 480
Db 653 CCTCAGAAATGTAAGCTTTACAGAGGGTCCGATCTGTTCCCAAGCAAGAGAG 712

QY 481 GCACATTTGAGTAATAGCTCCAAATTTGCACTGTTGCTGGCAACACAAATTTGAC 540
DB 713 AAACACCTGATGTAAGTCTGTTGACATCTGCAATCCCTTCTGCAAGGACGACGCTGAT 772
QY 541 TATGCAATTTTCTATGTTGAGCAAGTGGCGAATCAGAGTTTAAATCGCGGAAATTAATG 600
DB 773 TATGGCATCTAGCTCATCCACGAGGCTGAAGTAAAGTTTAAATGAGCAAACTCTTG 832
QY 601 AACGTTGATACGACGTAAGTACAGGCTTACCCATGAGAGTCTTCACTTTTCAATGAT 660
DB 833 AATGTTGGCTATGTAGAAAGCCCTCAAGGAAATTTGGAGCTTTTAAATTCACGAT 892
QY 661 GTGATTTTACTGCGCGAAGTGAACCGTAACGCTGTAACGCTCAATTCACCACTCAT 720
DB 893 GTGACCTGTTACCTGCAAGATGACTTTTACCTTTACAGTGTAGAGAGATCCCAAGCAT 952
QY 721 ATGAGTGTAGCGATCGATAAATTCATTAATTAATTCATATTTCCGATCTTCCGCGA 780
DB 953 CTGTTGTTGGCAGGAACAGCACTGGGTACAGTTACGTTACAGTGTATTTTGGGGGT 1012
QY 781 ATGAGTGTACATAAAGATCACTGTAAGAAATCAATGATTTTTCGATTTTTCG 840
DB 1013 GTTACTGCTTAAAGCAGAGAGAGTTTTCAGGTGATGATTTCTTAACTACTG 1072
QY 841 GGTGTTGGGCGAGAGAGCAGCATTTTGGCAGCAATCGATGCTGAGCTGAAAGTT 900
DB 1073 GATGTTGGGAGCGAAGCAGTACCTGACCTGAGGTTGAGCTCCCAAGATGAAATTT 1132
QY 901 TCAAGATATCCGACACAAATTTGACGATTAATTAATGATTAAGCATTCGACGAGCGACG 960
DB 1133 TCCCGGCCCCCTGCTGAAGTGGTAAATATACATGATGTTCTTCACTAGACAAAGGC 1192
QY 961 AATCCAGTTAATAATGCCCTCAAAATTAATGGCCCAAGCGCCGATGACGAT 1020
DB 1193 AATGAGGTGAACGAGACGATGAAGCTCTTACACCAAGTGTACAGATCTGAGAAC 1252
QY 1021 GACGCTTAAGCAATCTGAAGTAAAGCTGTAATCTGTAATCTGAAATGAACTCTTCACT 1080
DB 1253 GATGGGTGATGATGTTGTTCTTAATTAATGATGATCTGTGAGACAAATCTTTAATATC 1312
QY 1081 CGAGCGGTGTCGATTT 1097
DB 1313 AACATCACAGTGATTT 1329

RESULT 12
US-09-055-097-4
Sequence 4, Application US/09055097
Patent No. 5955282
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Puryi
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/055,097
FILING DATE: Filed Herewith

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0490 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNG10T06
CLONE: 2551161
US-09-055-097-4

Query Match 13.5%; Score 155.4; DB 2; Length 2280;
Best Local Similarity 51.8%; Pred. No. 7,6e-40; Mismatches 326; Indels 0; Gaps 0;
Matches 351; Conservative 0;
QY 421 CCTAAGATTTGTTTCAAGGACATGTTGCTATTAATGTCCTTAATAGATCGTGA 480
DB 771 CTTGAGAAATGTAAGCTTTTACAGAGGTCGCTATCTGTTCCCAAGGAGAGAG 830
QY 481 GCACATTTGAGTAATAGCTCCAAATTTGCACTGTTCTGCGCAACAAATTTGAC 540
DB 831 AAACACCTGATGTAAGCTGTTGAAATCTGCAATCTCTTCTGACAGGACGCTGAT 890
QY 541 TATGCAATTTTCTATGTTGAGAGAGAGGCGAATCAGACCTTTAATGCGGAAATTAAG 600
DB 891 TATGGATCTTACCTTACCAAGGCTGAAGTAAAGTTTAAATGAGCCTTAACTCTTG 950
QY 601 AACGTTGATTAAGCACTAGCATCAAGCTCTTCAATGCGAGCTTCACTTTTCATGAT 660
DB 951 AATGTTGGCTATTTAGAAAGCCCTCAAGGAAAGAAATTTGGAGCTGTTAATTTCCAGAT 1010
QY 661 GTGATTTTACTGCGCGAAGTGAACGTAACCTGTACAGCTGTCCAAATTCACCACTCAT 720
DB 1011 GTGACCTGTTACCGAGAAATGACTTTAATCTTTACAGTGTAGAGATCCCAAGCAT 1070
QY 721 ATGAGTGTAGCGATCGATAAATTCATTAATTAATCTTCAATTTGCGCGATCTTCCGCGA 780
DB 1071 CTGTTGTTGGCAGGAACAGCACTGGGTACAGTTACGTTACAGTGAATTTTGGGGGT 1130
QY 781 ATGAGTGTACATAAAGATCACTGTAAGAAATCAATGATTTTTCGATGATTTTTCG 840
DB 1131 GTTACTGCTTAAAGCAGAGAGAGGTTTTCAGGTGAATGATTTCTTAACTACTG 1190
QY 841 GGTGTTGGGCGAGAGAGCAGATTTTGGCGACGAGAAATCGATGCTGAGCTGAAGTT 900
DB 1191 GATGTTGGGAGGCGAAGACGATGACTCAGACTGAGGTTGAGCTCCAAAGAAATTAAT 1250
QY 901 TCAAGATATCCGACACAAATTTGACAGATTAATTAATTAAGTAAAGCTGACGAGAGGAG 960
DB 1251 TCCCGGCCCCCTGCTGAAGTGGTAAATATACATGATCTTCCACTGAGCAAAAGGC 1310
QY 961 AATCCAGTTAATAATGCGCTCAAAATTAATGAGGCAAGAGCGCGATGAGACAT 1020
DB 1311 AATGAGGTGAACGAGAGGATGAAGCTTTTACACCAAGTGTACAGAGCTGTGAGAAC 1370
QY 1021 GACGCTTAAGCAATCTGAAGTAAAGCTGTAATCTGGAATGGAAGCTCTTCACT 1080
DB 1371 GATGGTTGATGATGTTGTTCTTAATTAATGATGATCTGTGGAACAAATCTTTAATATC 1430
QY 1081 CGAGCGGTGTCGATTT 1097
DB 1431 AACATCACAGTGATTT 1447

RESULT 13

US-09-373-902-4

Sequence 4, Application US/09373902

Patent No. 6649737

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

Guegler, Karl J.

Corley, Neil C.

Shah, Puri

Paterson, Chandra

TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/373,902

FILING DATE: 12-Aug-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION NUMBER: US/09/055,097

APPLICATION DATA:

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Cerrone, Michael C.

REGISTRATION NUMBER: 39,132

REFERENCE/DOCKET NUMBER: PF-0490 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 2280 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: LUNGUT06

CLONE: 2551161

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-373-902-4

Query Match 13.5%; Score 155.4; DB 3; Length 2280;

Best Local Similarity 51.8%; Pred. No. 7.6e-40;

Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTGTCAGAGCATGCTGTGCTATTATGTCCTATAGAGATGTGA 480
DB 771 CCGCAGAAATGTAAGCTTTTACAGAGGTCGCGCATCTGTTCCCAACGGAACGAGG 830
QY 481 GCACATTTGAGATATATGCTCCACAAATTTGCACTGCTGCTGCCAACAACAAATTTGAC 540
DB 831 AAACACCTGATGTACCTGCTGAGACATCTGCATCCCTTCTGCAAGAGGACGCTGAT 890
QY 541 TATGCAATTTTCAATGCTGAGCAAGTGGCAATCAGACGTTTAATGCGGGAATCTAATG 600
DB 891 TATGCAATCTACGATCTCCACAGGCTGAAGGTAAGTTAATCGAACCAACTCTTG 950
QY 601 AAGCTTGATAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 951 AATGTTGGCTATCTAAGACCTTCAAGAAAGAAATTTGGAGCTGCTTTATATTTCCAGAT 1010
QY 661 GTGCAATTTACTGCCGAGATGACCGTAACTGTACAGCTGTCAATTCACACGCTCAT 720

DB 1011 GTGCACTGTATCCCGAGATGATCTTTAACCCTTACAGTGTGAGAGCATCCAGCAT 1070

QY 721 ATGAGTGTAGCATGTATTAATTTCAATTAATTAATTAATTAATTAATTAATTAATTAAT 780

DB 1071 CTGGTGTGTGCGAGAGACAGACCTGGGTACAGTTACGTTACAGTATATTTTGGGGGT 1130

QY 781 ATCAGTGTACTAACAAGATCACTGTAAGAAATCAATGATTTTCCAAATGATTTTGG 840

DB 1131 GTTACTGCTCCCTAGAGAGAGACGTTTTCAGAGGATGATTTCTTACAACTACTG 1190

QY 841 GGTGGGCGGAGAGACGACGATTTGGCGACGAGAACATGATGCTGAGCTGAAAGTT 900

DB 1191 GGATGGGAGGCGAAGAGATGACCTCAGACTCAGGGTTGAGCTCCAAGATGAAATTT 1250

QY 901 TCAATATTCGACACCAATTTGACGATTAATTAATTAATTAATTAATTAATTAATTAAT 960

DB 1251 TCCCGCCCTGCTGAAAGTGTAAATATTAATTAATTAATTAATTAATTAATTAATTAAT 1310

QY 961 AATCCAGTTAATTAATGCGCTACCAAAATTAATGCGCAAGAGCGCGATGACAGT 1020

DB 1311 AATGAGTGAACGAGAACGATGAGCTTTACACCAAGTGTACAGAGTCTGAGAGCA 1370

QY 1021 GACGCTTAAGCATCTGATGATTAAGCTGTAATTTGGAATTTGAGCTCTTACACT 1080

DB 1371 GATGGTTGAGTATGTTCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1430

QY 1081 CGAGCGCTGCTGATTT 1097

DB 1431 AACATCAGAGTGATTT 1447

RESULT 14

US-09-949-016-4375

Sequence 4375, Application US/09949016

Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

CURRENT FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4375

LENGTH: 2965

TYPE: DNA

ORGANISM: Human

US-09-949-016-4375

Query Match 13.0%; Score 150; DB 3; Length 2965;

Best Local Similarity 56.8%; Pred. No. 5.1e-38;

Matches 276; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

QY 405 CGGTGACATGGAATGCTTAAGATTGTGTCAGAGCATGCTGTGCTATTATTTGTC 464
DB 336 CGAGGTACTGGAAGCTTTGATTTGATGCTCTGGTGGAAAGTGGGATCTTATTC 395
QY 465 CTATAGAGTGTGGAAGACATTTGAGATTAATGCTCCACAAATTTGCACTGCTGCTGC 524
DB 396 CTTCGGAACGCGACGACACTCCAGTCTGTTTGAAGACACTGCTTCCATCTCCA 455
QY 525 CAACAAATTTGATCTATGATTTTATTTGATGAGGAAGTGGGATTCAGAGCTTAA 584
DB 456 GCGCAGGCGCTGCACTTTGATTTATGTTGTAACAAGTTGGTATCCCAACCTTTAA 515

XX Claim 5; Fig 1; 51bp; English.

CC The present sequence encodes a beta1,4-N-acetylgalactosaminyltransferase,
 CC designated beta4GalNAcT. The enzyme is required for the biosynthesis of
 CC animal cell glycoproteins. The enzyme functions to synthesize the
 CC LacdINac or LDN motif GalNAcbeta4GlcNAc-R. This motif is glycoprotein
 CC hormones produced by the pituitary gland, and is a major marker of
 CC glycoconjugates made by parasitic and non-parasitic invertebrates and may
 CC be implicated in host immune responses to infection. The protein and
 CC polynucleotides are useful for transforming or transfecting host cells
 CC for producing substantially pure forms of the enzyme, or in vitro, for
 CC formation of a LDN structure on proteins or peptides.

CC Sequence 1152 BP; 314 A; 275 C; 259 G; 304 T; 0 U; 0 Other;

Query Match 100.0%; Score 1152; DB 12; Length 1152;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCTTTTCGATTTGGAGTCGCGCAGACTCAAGTCGTTGCTCGACTTTGTGCGCTT 60
 Db 1 ATGGCTTTTCGATTTGGAGTCGCGCAGACTCAAGTCGTTGCTCGACTTTGTGCGCTT 60
 QY 61 CTTCTATTAGTTCATGCAATGATTTTAAAGATTCACATCGCTTTACGAGAACCTTACTATC 120
 Db 61 CTTCTATTAGTTCATGCAATGATTTTAAAGATTCACATCGCTTTACGAGAACCTTACTATC 120
 QY 121 GGGTCTCTCGACCTTATTTGCGCAGTCGACGCAATGAGAGCAATGCTCGGGAATACGGCT 180
 Db 121 GGGTCTCTCGACCTTATTTGCGCAGTCGACGCAATGAGAGCAATGCTCGGGAATACGGCT 180
 QY 181 TCCACTTCGGATGATCTACTTGTATGATGAGATTTCCAGTTTCCACGATTTCTGAAGTT 240
 Db 181 TCCACTTCGGATGATCTACTTGTATGATGAGATTTCCAGTTTCCACGATTTCTGAAGTT 240
 QY 241 AATCAGACTAGTTTATGAGAGCAATTCGTCAATCTGTTCCCGACCAACAGACTCTT 300
 Db 241 AATCAGACTAGTTTATGAGAGCAATTCGTCAATCTGTTCCCGACCAACAGACTCTT 300
 QY 301 CAATTCTGTATCAGACACTCCCACTGCTGGAGCCATCCGTTGATTTCTTCGATGAG 360
 Db 301 CAATTCTGTATCAGACACTCCCACTGCTGGAGCCATCCGTTGATTTCTTCGATGAG 360
 QY 361 CCCGACTTCAAACTCTCGAAGAAATCTATCCGAGACGACGCGGTGACATGGAATG 420
 Db 361 CCCGACTTCAAACTCTCGAAGAAATCTATCCGAGACGACGCGGTGACATGGAATG 420
 QY 421 CCTAAGGATTTGTGTGCAAGGCAATCGTGTGCTATTATTTGTCCTTATAGATCGTGA 480
 Db 421 CCTAAGGATTTGTGTGCAAGGCAATCGTGTGCTATTATTTGTCCTTATAGATCGTGA 480
 QY 481 GCACATTTGGAATATGCTCCACAAATTTTGCACTCGTTGCTCGCAAAACAATTTGAC 540
 Db 481 GCACATTTGGAATATGCTCCACAAATTTTGCACTCGTTGCTCGCAAAACAATTTGAC 540
 QY 541 TATGCAATTTTCATTTGGAAGCAAGTGGCGAATGAGAGCTTTTATGCGGGAAACTAATG 600
 Db 541 TATGCAATTTTCATTTGGAAGCAAGTGGCGAATGAGAGCTTTTATGCGGGAAACTAATG 600
 QY 601 AACGTTGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 Db 601 AACGTTGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 661 GTGCAATTTACTGCGCGAAGATGACCGTACCTGTACACGTTGCTCAATTTCAACACGTCAT 720
 Db 661 GTGCAATTTACTGCGCGAAGATGACCGTACCTGTACACGTTGCTCAATTTCAACACGTCAT 720
 QY 721 ATAGAGTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 Db 721 ATAGAGTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 QY 781 ATCAGTGCACCTAACAAAAGATCACTGTAAGAAATCAATGATTTTGGAAATGATTTTGG 840

Db 781 ATCAGTGCACCTAACAAAAGATCACTGTAAGAAATCAATGATTTTGGAAATGATTTTGG 840
 QY 841 GGTGTTGGGGGCGAGAGAGAGAGATTTGGCGACAGCAATCATGATGCTGGAAGTT 900
 Db 841 GGTGTTGGGGGCGAGAGAGAGAGATTTGGCGACAGCAATCATGATGCTGGAAGTT 900
 QY 901 TCAAGATATCCGACACAAATTTGCAAGATATTAATGATTTAGCACTGACGGAAGCGAG 960
 Db 901 TCAAGATATCCGACACAAATTTGCAAGATATTAATGATTTAGCACTGACGGAAGCGAG 960
 QY 961 AATCCAGTTAATTAATGCGGCTTACAAATATATGCGCAACGAGAGCGGATGACAGCT 1020
 Db 961 AATCCAGTTAATTAATGCGGCTTACAAATATATGCGGCAACGAGAGCGGATGACAGCT 1020
 QY 1021 GACGGCTTAAGCAATCTGAAGTATAGCTCCGTAATCTGGAATTTGAAGCTCTTACACT 1080
 Db 1021 GACGGCTTAAGCAATCTGAAGTATAGCTCCGTAATCTGGAATTTGAAGCTCTTACACT 1080
 QY 1081 CGAGCCGTCGTGATTTTGTCTGAAAAGAGACTCCGCGGAGCTGCAAGGAGCTTTCCA 1140
 Db 1081 CGAGCCGTCGTGATTTTGTCTGAAAAGAGACTCCGCGGAGCTGCAAGGAGCTTTCCA 1140
 QY 1141 ACGTGTTTTAA 1152
 Db 1141 ACGTGTTTTAA 1152

RESULT 2
 ABL14209
 ID ABL14209 standard; cDNA, 1839 BP.
 XX
 XX ABL14209;
 DE 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 37109.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical; gene; ss.
 XX
 XX Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 XX
 XX 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PMD, Myers EW;
 XX
 XX WPI; 2001-656860/75.
 XX P-P-SDB; ABB70106.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 XX genes from Drosophila and for elucidating cell signalling and cell-cell
 XX interactions.
 XX
 XX Claim 1; SEQ ID NO 37109; 21bp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (ABL16176-ABJ30511), expressed DNA
 XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 XX ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly
 CC from WFO at ftp.wfo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1839 BP; 501 A; 450 C; 384 G; 504 T; 0 U; 0 Other;

Query Match 16.7%; Score 192.6; DB 4; Length 1839;
 Best Local Similarity 58.7%; Pred. No. 9.9e-52;
 Matches 333; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

399 GCACGCCGCTGACATGGAATGCTTAAGATTGTTGAGAGGATCGGTGTTATTAT 458
 DB GCGCCCTGATGCGCTTCGACGCTGAAATGCAATGCGCATGCGCATGCTGCTATTGT 659
 459 TGTGCGCTATAGATGCTGAGACGATTTGAAATGCTCCCAATTTGCTGCTGTT 518
 DB TGTGCGCTATGCGCATGCTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 719
 519 GCTGCGCAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 578
 DB TCTGATGAGAGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 779
 579 GTTTAATGCGGGAACATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 638
 DB CTTTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
 639 GCAAGCTCTTCATCTTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698
 DB GATGCTGCTTTCATCTTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
 840 GATGCTGCTTTCATCTTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 899
 699 GTGTCATTCATTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758
 DB CTGTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
 900 CTGTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959
 759 ATATTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 818
 DB TATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
 960 TATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1019
 819 TGTGATTCGAAATGATTTTGGGGTGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 878
 DB TGGGTTTCGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
 1020 TGGGTTTCGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1079
 879 ATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 938
 DB GAAGCAGCGAATCTATTCATATCAAGGATCCGGTCAATAGCTCGCTCAAGATGCT 1139
 1080 GAAGCAGCGAATCTATTCATATCAAGGATCCGGTCAATAGCTCGCTCAAGATGCT 1139
 939 TAAGCATCTGACGAG 965
 DB GAAAGCATGAG 1166
 1140 GAAAGCATGAG

RESULT 3
 AA242120
 ID AA242120 standard; cDNA; 1938 BP.

AA242120;
 31-JAN-2000 (first entry)
 Human endometrium tumour cDNA derived EST 140.
 Endometrium; human; tumour; cancer; anticancer; cytostatic;
 EST: treatment; uterine; gene therapy; expressed sequence tag; ss.

OS Homo sapiens.
 XX
 XX DE19817948-A1.
 XX
 XX 21-OCT-1999.
 XX
 XX 17-APR-1998; 98DE-01017948.
 XX
 XX 17-APR-1998; 98DE-01017948.
 XX

PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 XX WPI; 1999-591957/51.
 DR
 XX
 XX
 PT New nucleic acid sequences expressed in uterine cancer tissues, and
 PT derived polypeptides, for treatment of uterine and endometrial cancer and
 PT identification of therapeutic agents.
 PS Claim 3; Page 273; 444dp; German.

This invention describes novel human nucleic acid (cDNA) sequences (A),
 CC that are highly expressed in uterine tumour tissue and which have
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
 CC are used (i) to identify agents suitable for treatment of uterine or
 CC endometrial cancer; (ii) directly for treating these forms of cancer
 CC (including expression from gene therapy vectors) and (iii) for generation
 CC of specific antibodies. (A) are identified by assembling ESTs (expressed
 CC sequence tags) from a particular tissue type before comparison of
 CC expression patterns. This allows a significantly longer fragment of the
 CC gene to be revealed, so should reduce the number of failures associated
 CC with the fact that ESTs from different libraries may represent different
 CC parts of the same unknown gene, distorting the estimated frequency of
 CC occurrence in a particular tissue. AA241981-242121 represent EST
 CC fragments derived from a human endometrium tumour cDNA library which
 CC encode the protein sequences represented in AA59991-760328
 XX

Sequence 1938 BP; 395 A; 567 C; 515 G; 461 T; 0 U; 0 Other;

Query Match 14.5%; Score 167.6; DB 2; Length 1938;
 Best Local Similarity 54.4%; Pred. No. 1.7e-43;
 Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

421 CCTAAGATGTTGTTGCAAGGAGCTGTTGCTATTTATTTGCTTATAGAGATGTA 480
 DB CTTGAGGTTGTTGAGCCCGCTCCGGAACGACATGTTGCTCATGCTGCTGCTGCTGCTGCT 632
 573 CTTGAGGTTGTTGAGCCCGCTCCGGAACGACATGTTGCTCATGCTGCTGCTGCTGCTGCT 632
 461 GCACATTTGAAATATGCTCACAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 DB CACCACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
 633 CACCACTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
 541 TATGCAATTTTCATGTTGAG 600
 DB TATGCAATTTTCATGTTGAG 752
 693 TATGCAATTTTCATGTTGAG 752
 601 AACGTTGATGAG 660
 DB AACGTTGAG 812
 753 AACGTTGAG 812
 661 GTGATTTTACTGCGGAG 717
 DB GTGATTTTACTGCGGAG 872
 813 GTGATTTTACTGCGGAG 872
 CATATGAGGTAG 777
 DB CATATGAGGTAG 932
 873 CATATGAGGTAG 932
 778 GGAATCAGTGCACTAACAAG 837
 DB GGAATCAGTGCACTAACAAG 992
 993 GGAATCAGTGCACTAACAAG 992
 838 TGGGTTTGGGCGAG 897
 DB TGGGTTTGGGCGAG 1052
 993 TGGGTTTGGGCGAG 1052
 898 GTTTCAGATATCCGACCAATTCAGATATTAATGATTAATGATTAATGATTAATGATTAATG 957
 DB GTTTCAGATATCCGACCAATTCAGATATTAATGATTAATGATTAATGATTAATGATTAATG 1112
 1053 ATCTTCGCGCCCGACATCTGTAGAGACATTAATGATTAATGATTAATGATTAATGATTAATG 1112
 958 ACGAATTCAGTTAATTAATGCGGCTACAAATTAATGCGGCTACAAATTAATGCGGCTACGAG 1017

DB 1113 GGCAATGAGAAAATCCACAGATTGACCTTCGTCGCTACCCAGAAATTCCTGAGAC 1172
QY 1018 CGTGACGGCCTTAAGCATCTGAAGTAACTGTAATCTGAAGTGAAGCCTCTCTAC 1077
DB 1173 CAAGATGGAGATGAATCTCAGTACATCACTGCTGCTGAGAGCTGGGGCTCTTTAT 1232
QY 1078 AC 1079
DB 1233 AC 1234
RESULT 4
ACN37523
ID ACN37523 standard; cDNA; 1949 BP.
XX ACN37523;
XX ACN37523;
XX 18-NOV-2004 (first entry)
XX Tumour-associated antigenic target (TAT) cDNA DNA323958, SEQ ID NO:468.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX Tumour; diagnosis; cell proliferative disorder; breast cancer;
XX colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX central nervous system cancer; bladder cancer; pancreatic cancer;
XX cervical cancer; melanoma; leukemia; hybridisation probe;
XX chromosome identification; chromosome mapping; gene mapping;
XX gene therapy; cytostatic; gene; ss.
XX
XX Homo sapiens.
XX
XX WO2004030615-A2.
XX
XX 15-APR-2004.
XX
XX 29-SEP-2003; 2003WO-US028547.
XX
XX 02-OCT-2002; 2002US-0414971P.
XX
XX (GETH) GENENTECH INC.
XX Wu TD, Zhang Z, Zhou Y;
XX
XX MPI; 2004-347921/32.
XX P-PsDB; ABM80189.
XX
XX New tumor-associated antigenic target polypeptides and nucleic acids,
XX useful in preparing a medicament for treating or detecting a
XX proliferative disorder, e.g. breast, lung, colorectal, ovarian or
XX prostate cancer or tumor.
XX
XX Claim 1; SEQ ID NO 468; 7273pp; English.
XX
XX The invention relates to human tumour-associated antigenic target (TAT)
XX polypeptides, and their related nucleic acids. The TAT polypeptides are
XX overexpressed in cancer tissues compared to normal tissues, and may thus
XX serve as effective targets for the diagnosis and treatment of cancer in
XX mammals. The invention also relates to nucleic acid and polypeptide
XX sequences at least 80% identical to the TAT nucleic acids and
XX polypeptides; expression vectors and host cells comprising a TAT nucleic
XX acid; an antibody specific for a TAT polypeptide; a peptide or organic
XX molecule which binds to a TAT polypeptide; fusion proteins comprising a
XX TAT polypeptide; and methods and compositions for the treatment or
XX diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
XX antibodies, antagonists, binding molecules and compositions are useful
XX for diagnosing or treating a cell proliferative disorder associated with
XX increased TAT expression, particularly cancers such as breast cancer,
XX colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
XX cancer, pancreatic cancer, cervical cancer, cancers of the central
XX nervous system, melanoma and leukemia. TAT nucleic acids may further be
XX used as hybridisation probes, in chromosome and gene mapping, in
XX chromosome identification and in gene therapy. The present sequence
XX represents a TAT nucleic acid of the invention

XX SQ Sequence 1949 BP; 376 A; 584 C; 525 G; 464 T; 0 U; 0 Other;
Query Match 14.5%; Score 167.6; DB 13; Length 1949;
Best Local Similarity 54.4%; Pred. No. 1.7e-43;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;
QY 421 CCTAAGATTTGTGTGCAAGGCGATCGTGTGCTATTTATGTCCTATAGAGATGTGAA 480
DB CCTGCAGGTTGTGAGCCCGCGCTCCGAAACAGCATATGTGCTCATGTGCCGGGAG 667
QY 481 GCACATTTGAGATATATGCTCCACAATTTGCACTGTGCTGCCAACAACATTTGAGC 540
DB CACCACTGCGGCTGCTGCTTACCACTGCACCCCTTCTTACAGCGCAGAGAGCTTGT 727
QY 541 TATGCAATTTTCATTGTGAGAGCAAGGCGCATCAAGCTTTAATTCGCGGAAACATGAT 600
DB TATGGCATTTATGTATCTATCCACAGAGCTGGAATTTGAACATTTAACAAGGCAAACTGTTG 787
QY 601 AACGTTGATGACGAGTACATCAGCCTCTTACCCATGCGAGTGTTCATCTTTCATGAT 660
DB AAGTTGGGGGTGCGAGAGGCCCTGGGTGATGAAGAGTGGAGCTGCTGTTTGACGAT 847
QY 661 GTGATTTAATGCGCGAGATGACCGTAACTGTACAGTGT---CCAATTCAACGACGT 717
DB GTGACCTTCTGCGAAGAAATGACCAATCTGTATGTGTGAGACCCCGGGGACCCCGC 907
QY 718 CATATGAGTGTACGATGATGATTAATTCATTTAATTCATATTCGCGATCTTGGC 777
DB CAGTTGCGGTTGCTATGAAACAAAGTTTGATACAGCTTCCGTAACCCAGTACTTTCGA 967
QY 778 GGAATCAGTCACTAACAAAGATCACTGAAGAAATCAATGATTTTGAATGATTTT 837
DB GGAATCTCAGCACTTAATCTCTGACGATCACTGAAGATGAATGGCTTCCCATGAAATAC 1027
QY 838 TGGGGTTGGGGCGAGAGAGACGATTTTGGGACGAGAAACATCATGCTGTGACCTGAAA 897
DB TGGGGCTGGGGGTGTGAGATGACGATTTGCTACCAAGAGTGGCGCTGTGGGATGAAAG 1087
QY 898 GTTTCAGATATTCGACCAAAATGACAGATTAATTAATTAAGCACTGCAGGAAGCG 957
DB ATCTTCGGCCCCCAGATCTGTAGGACCTATTAATGTATGAGACCGAGAGATTAAG 1147
QY 958 ACGAATCGATTAAATATGCGCTACAAATTAATGAGCCAAAGAGCCGATGAGACA 1017
DB GGCAATGAGAGAAATCCACAGATTGACCTTCGTCGCTGACCAAGATTCCTGAGACG 1207
QY 1018 CGTGACGGCCTTAAGCATCTGAAGTAACTGTAATCTGTAATCTGAATGAAAGCTCTCTAC 1077
DB 1208 CAAGATGGAGATGAATCACTGACATCACTGCTGCTGAGAGCTGGGGCTCTTTAT 1267
QY 1078 AC 1079
DB 1268 AC 1269
RESULT 5
AAC77836
ID AAC77836 standard; cDNA; 2002 BP.
XX AAC77836;
XX
XX 08-FEB-2001 (first entry)
XX
XX Human cancer associated gene sequence SEQ ID NO:230.
XX
XX Human; cancer associated gene; cancer antigen; detection; cancer;
XX diagnosis; cytostatic; proliferative; vulnerability; immunomodulator;
XX antidiabetic; antihypertensive; antineoplastic; antibacterial; antiviral;
XX anti-inflammatory; antihypertensive; antidiabetic; antibacterial; anti-
XX dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
XX vasodilator; antihypertensive; angiogenic; gene therapy; inflammation;
XX immune disorder; haematopoietic cell disorder; autoimmune disorder;

PT New isolated NOVX polypeptides and polynucleotides, useful for
 PT preventing, diagnosing or treating NOVX-associated disorders, e.g.
 PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
 PT asthma, or infections.

XX Claim 20; SEQ ID NO 287; 800pp; English.

CC The invention relates to novel isolated polypeptides, mature forms of
 CC these, or a sequence that is at least 95 % identical to, or having one or
 CC more conservative amino acid substitutions in the polypeptides. The
 CC polypeptides, nucleic acid molecules and antibodies are useful in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease, preferably a NOVX-associated disorder. The nucleic acid
 CC molecules, polypeptides and antibodies are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
 CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
 CC atherosclerosis), neurodegenerative disorders, Alzheimer's disease,
 CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
 CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
 CC dyslipidemias. The nucleic acids and polypeptides may also be used as
 CC targets for the identification of small molecules that modulate or
 CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
 CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in
 CC generation of antibodies that bind immunospecifically to NOVX substances
 CC for use in therapeutic or diagnostic methods. The nucleic acids are
 CC further used as hybridization probes, in chromosome mapping, tissue
 CC typing, preventive medicine, and pharmacogenomics. This sequence
 CC corresponds to the gene encoding one of the NOVX polypeptides of the
 CC invention.

XX Sequence 1062 BP; 303 A; 224 C; 231 G; 304 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 10; Length 1062;
 Best Local Similarity 54.3%; Pred. No. 5.1e-41;
 Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACACGACCGCGTGACATGATGCTTAAGATTGTGTGCAAGCATCTG 448
 DB 297 ATTGATATTTAGAGCGAGGGGTCATTGAGCGCAAAACCTGTAACCACTGGAAG 356
 QY 449 TTGCTATTATTTGTCCTTATAGATCGTGAACGACATTGAGATATATCTCCACATT 508
 DB 357 TGGCAGTTCTCATCTTCCGTAATCGCATGACATCTCCAAATTTTCTTACATC 416
 QY 509 TGCATCTGTTGTCGCGCAACCAATTTGACATATCTTCAATTTGAGCAAGTGG 568
 DB 417 TCAATTCAGATGCTCCGAAAGACGCGCTGAATTTGCTTTATATGTCATTGAACACTG 476
 QY 569 CGAATCAGAGCTTTAATCGCGGAACTATGAACTTGGATACGAGTACATCAGCC 628
 DB 477 GCACACACCTTTAATCGTGCATGCTTTCAATGTGGCTTCAAAAGGCCATGAAG 536
 QY 629 TCTACCATGCGACGCTTCAATCTTTCATGATGTCGATTAACTGCCGAAGTACCGTA 688
 DB 537 ACAGTGTCTGGACCTGTATATCTTCCAGATGTGATCATCTCAAAATGACCGGA 596
 QY 689 ACCTGACAGCTGTCATTAATCAACCGCTCATATAGTGTAGAGATCGATAATTTCAAT 748
 DB 597 ACTATTAACGATGTGGAAGATGCAACGCTATTTGCTCAAAAGCTGATTAATACATGT 656
 QY 749 ATAAATCTCATATTTGCGGATCTTGGCGAATCAGTCACTAACAAAGATCACTGA 808
 DB 657 ATATTTCTCATATAAAGATTTTGTGTGTATAGTGGCTGACAGTGTGAACAATTTA 716
 QY 809 AGAAAAATCAATGATTTTTCGATGATTTTGGGTTGGGGCGAGAGACGACGATTTGG 868
 DB 717 GAAAGATCATATGTTTCTTAATGCTTCTGGGAGTGGGAGGAGATGATGACCTTT 776
 QY 869 CGACGAGACATCGATGCTGACCTGAAGTTTCAAGATATCCGACAAATTTGACGAT 928
 DB 777 GGAACAGATTCTATATGCTGTGATATATGTAACGACCAAGAGGAGACTTGAAGAAAT 836

QY 929 ATAAATGATTAAGACTCGACGGAAGCGAATTCAGTTATATAATCCGCTACAAA 988
 DB 837 ACAAGTCAATTTCTCTCA-TTACCATAGAGGTGAAGTCACTT--TTAGACCGGTATAAAT 893
 QY 989 TATGCGCCAAAGAAAGCCCGCATGACACGTGACGCTTAAGCAATGGAATTAAC 1048
 DB 894 TACTAAGGTATTCCAAGGACGTCAGTACATGATGAGTGAACAATTTATATATAGGC 953
 QY 1049 TCGTAATCTGG 1060
 DB 954 CAAAATTAAGTGG 965

RESULT 8

AAV69749 standard; cDNA; 1146 BP.

AAV69749;

04-FEB-1999 (first entry)

Human lactosyl ceramide synthase encoding cDNA.

Lactosyl ceramide synthase; enzyme; recombinant; rat; mouse; human;
 uridine phosphate-galactose; glycosyl ceramide; sphingoglycolipid; ds.

Homo sapiens.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Location/Qualifiers
 1. 1146
 /*tag= a
 /product= "Lactosyl ceramide synthase"
 /note= "the stop codon is not indicated"

JP1029571-A.

10-NOV-1998.

25-FEB-1998; 98UP-00043335.

28-FEB-1997; 97JP-00045401.

(KAOS) KAO CORP.

(RIKA) RIKAKU KENKYUSHO.

WPI, 1999-038272/04.

P-PSDB; AAW81569.

Lactosyl ceramide synthase and its gene - used for recombinant production of the enzyme.

Claim 4; Page 13; 15pp; Japanese.

This cDNA encodes a human lactosyl ceramide synthase. The invention provides rat, mouse and human lactosyl ceramide synthases (AAW81567 to AAW81569) and genes (AAV69747 to AAV69749) encoding the enzymes respectively. The nucleic acids, and host cells transformed with a recombinant DNA containing the lactosyl ceramide synthase gene are used for the recombinant production of the enzyme. The new lactosyl ceramide synthase has the following enzymatic properties: it transfers galactose from uridine phosphate-galactose to glycosyl ceramide to synthesize lactosyl ceramide; it reacts strongly with glycosyl ceramide and weakly with globoside and does not substantially with ceramide, lactosyl ceramide, GA2 and GM2. Lactosyl ceramide is used as a raw material for sphingoglycolipid

Sequence 1146 BP; 339 A; 240 C; 242 G; 325 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 2; Length 1146;
 Best Local Similarity 54.3%; Pred. No. 5.3e-41;
 Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

389 ATCCGACACGACCGCGTGACATGAAATGCTTAAGATTGTGTGCAAGCATCTG 448

```

Db 413 ATTTAGATATTGAGCAGGGGGTCATTGGAGGCGAAGAGCTGTAAACCAATGGAAG 472
Qy 449 TTGCTATTATTGTCGCCCTATAGATCGTAGACCAATTGGAGATTAAAGCTCCAAATT 508
Db 473 TGGCAGTCTTCATCTCTTCCTGTAATGCGCATMAACATCTTCCAAATTTTCTTACATC 532
Qy 509 TGCACTCGTTGCTCGGCCAAACAAACAATTGACTATGCAATTTTCAATTGAGCAAGTGG 568
Db 533 TGAATTCAAATGCTCCAGAAAGACGGCTGGAATTTGCGTTTATGCTATTGAACACACTG 592
Qy 569 CGAATCAGACGTTTAAATCGCGGAAACATAAGACGTTGATACGACGATCAGCC 628
Db 593 GGAACAACCTTTTAAACGTCGATGCTTTTCATGTCGGCTTCAAGAGGCCATGAAG 652
Qy 629 TCTACCAATGACAGTCTTCATCTTTCATATGATCGATTTACTGCCCCGAAGATGACCGTA 688
Db 653 ACAGTCTCGGGACTGTGTAATCTTCCACGATGTGATCATCTACCTGAAGATGACCGGA 712
Qy 689 ACCGTACACGTGTCCAAATTCACACGATCATATGATGATGAGGATCGATTAATTCATT 748
Db 713 ACTATTACGATGTGAGAAATCCACGTCATTTTGTCTGCAAGCTGATTAATACATGT 772
Qy 749 ATAACTTCATATTGCGCGATCTTCGCGGAATCAGTGCATPAACAAAGATCACCTGA 808
Db 773 ATATTCTTCATATAAGAAATTTTGTGTGTGAAGTGGCTGACAGTGAACAATTA 832
Qy 809 AGAAATTCATGATTTTGCATGATTTTGGGGTGGGGCGAGAGACAGATTTGG 868
Db 833 GAAAGATCAATGTTTCTTAATGCTTCTGGGGATGGGAGAGAAATGATGACCTTT 892
Qy 869 CGACGAAACATGATGATGCTGACCTGAAAGTTTCAAGATATCCGACCAAAATGACGAT 928
Db 893 GGAACGAGTTCATCTATGCTGATATATATTAACACAGACAGAGGAGACTTGAAGAAAT 952
Qy 929 ATAAATGATTAAGCACTCGACGAAACGACGAATCCAGTTAAATGATGCGGTACAAA 988
Db 953 ACAAGTCAATTCCTCA-TCAACATAGAGTGAAGTCCAGTT--TTAGACGGTATTAAT 1009
Qy 989 TAATGGCCCAAGAGACCGCGATGACACGTCAGCGCTTAAGCAATCTGAATTAAGC 1048
Db 1010 TACTTAAGGTATTCAGAGAGCGTCAGTACATGATGAGTGAACAATTTAATATAGGC 1069
Qy 1049 TCGTAATCTGG 1060
Db 1070 CAAAATACTGG 1081

RESULT 9
ADJ95057 standard; DNA; 1201 BP.
AC ADJ95057;
AC 06-MAY-2004 (first entry)
DE Novel NOVX gene sequence #143.
XX
XX de; gene; antidiabetic; anorectic; cardiast; hypotensive;
XX antiarteriosclerotic; anorectic; virucide; antibacterial; fungicide;
XX antiparasitic; antiparasitic; antiparasitic; antiparasitic;
XX anticonvulsant; osteopathic; antihypertensive; antiparasitic;
XX dermatological; antihypertensive; antiparasitic; antiparasitic;
XX metabolic disorder; diabetes; obesity; infectious disease; anorexia;
XX cancer; cardiovascular disease; hypertension; atherosclerosis;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX epilepsy; immune disorder; osteoarthritis; hematopoietic disorder;
XX inflammatory skin disorder; asthma; dyslipidemia; neurogenesis;
XX cell differentiation; cell proliferation; hematopoiesis; wound healing;
XX angiogenesis; chromosome mapping; tissue typing; pharmacogenomic.
XX Homo sapiens.
XX

```

```

PN WO2003040325-A2.
XX
PD 15-MAY-2003.
XX
XX 05-NOV-2002; 2002MO-US035464.
XX
XX 05-NOV-2001; 2001US-0338626P.
XX 06-NOV-2001; 2001US-033072P.
XX 09-NOV-2001; 2001US-0348283P.
XX 15-NOV-2001; 2001US-0335610P.
XX 16-NOV-2001; 2001US-0338543P.
XX 20-NOV-2001; 2001US-0331630P.
XX 20-NOV-2001; 2001US-0331641P.
XX 21-NOV-2001; 2001US-0332152P.
XX 27-NOV-2001; 2001US-0333461P.
XX 28-NOV-2001; 2001US-0333912P.
XX 28-NOV-2001; 2001US-0334027P.
XX 29-NOV-2001; 2001US-0334300P.
XX 30-NOV-2001; 2001US-0334421P.
XX 30-NOV-2001; 2001US-0334526P.
XX 04-DEC-2001; 2001US-0336576P.
XX 04-DEC-2001; 2001US-033664P.
XX 07-DEC-2001; 2001US-0338314P.
XX 07-DEC-2001; 2001US-0338390P.
XX 10-DEC-2001; 2001US-0339008P.
XX 10-DEC-2001; 2001US-0339008P.
XX 11-DEC-2001; 2001US-0339286P.
XX 01-FEB-2002; 2002US-0353280P.
XX 01-FEB-2002; 2002US-0353288P.
XX 04-FEB-2002; 2002US-0354392P.
XX 04-FEB-2002; 2002US-0354393P.
XX 04-FEB-2002; 2002US-0354409P.
XX 27-FEB-2002; 2002US-0359944P.
XX 27-FEB-2002; 2002US-0360148P.
XX 05-MAR-2002; 2002US-0361790P.
XX 05-MAR-2002; 2002US-0361833P.
XX 05-MAR-2002; 2002US-0361925P.
XX 05-MAR-2002; 2002US-0362230P.
XX 13-MAR-2002; 2002US-0362625P.
XX 13-MAR-2002; 2002US-0364000P.
XX 13-MAR-2002; 2002US-0364181P.
XX 13-MAR-2002; 2002US-0364182P.
XX 13-MAR-2002; 2002US-0364197P.
XX 13-MAR-2002; 2002US-0364221P.
XX 17-MAY-2002; 2002US-0381621P.
XX 28-MAY-2002; 2002US-0383675P.
XX 17-JUL-2002; 2002US-0396703P.
XX 06-AUG-2002; 2002US-0401552P.
XX 07-AUG-2002; 2002US-0401594P.
XX 07-AUG-2002; 2002US-0401787P.
XX 15-AUG-2002; 2002US-040319P.
XX 20-AUG-2002; 2002US-0404821P.
XX 23-AUG-2002; 2002US-0405368P.
XX 23-AUG-2002; 2002US-0405402P.
XX 23-AUG-2002; 2002US-0405462P.
XX 23-AUG-2002; 2002US-0405631P.
XX 25-AUG-2002; 2002US-0406125P.
XX 04-NOV-2002; 2002US-00287226.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Agee ML, Alsbrook JP, Berghs C, Boldog FL, Burgess CE, Chant JS,
XX Chaudhuri A, Dipippo VA, Edinger SR, Eissen A, Ellerman K,
XX Gangoli EA, Gorman L, Gerlach VR, Ji W, Kekuda R, Khramsov NV,
XX Li L, Malynkar UM, MacDougall JR, Mezes PS, Miller CE, Millet I,
XX Ooi CE, Ort T, Padigar M, Patlurajan M, Raetelli L, Rieger DK,
XX Rothenberg ME, Shenoy SG, Spaderna SK, Spytek KA, Taupier RJ,
XX Verneer CM, Zerhusen BD, Zhong M;
XX
XX WPI; 2003-441551/41.
XX P-PSDB; ADJ95058.
XX
XX New isolated NOVX polypeptides and polynucleotides, useful for

```


PT preventing, diagnosing or treating NOX-associated disorders, e.g.
PT osteoarthritis, obesity, atherosclerosis, cancer, Parkinson's disease,
PT asthma, or infections.

PS Claim 20; SEQ ID NO 285; 800bp; English.

XX
XX
XX The invention relates to novel isolated polypeptides, mature forms of
CC these, or a sequence that is at least 95 % identical to, or having one or
CC more conservative amino acid substitutions in the polypeptides. The
CC polypeptides, nucleic acid molecules and antibodies are useful in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease, preferably a NOX-associated disorder. The nucleic acid
CC molecules, polypeptides and antibodies are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious diseases (viral, bacterial, fungal, helminthic, and
CC protozoal), anorexia, cancer, cardiovascular diseases (hypertension,
CC atherosclerosis), neurodegenerative disorders (Alzheimer's disease,
CC Parkinson's disease, epilepsy, immune disorders (osteoarthritis),
CC hematopoietic disorders, inflammatory skin disorders, asthma, and various
CC dyslipidemias. The nucleic acids and polypeptides may also be used as
CC targets for the identification of small molecules that modulate or
CC inhibit e.g. neurogenesis, cell differentiation, cell proliferation,
CC hematopoiesis, wound healing and angiogenesis, in gene therapy, in
CC generation of antibodies that bind immunospecifically to NOX substances
CC for use in therapeutic or diagnostic methods. The nucleic acids are
CC further used as hybridization probes, in chromosome mapping, tissue
CC typing, preventive medicine, and pharmacogenomics. This sequence
CC corresponds to the gene encoding one of the NOX polypeptides of the
CC invention.

XX
XX
XX Sequence 1201 BP; 355 A; 250 C; 262 G; 334 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 10; Length 1201;

Best Local Similarity 54.3%; Pred. No. 5.4e-41;

Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

QY 389 ATCCGACACGACGCGCGTGACATGATGCTTAAGATTGTGTGACAGCATCTGTG 448
DB 436 ATTTAGATATTGAGCGAGGGGTCATTGGAGGCCAAGAACCTGTAACCCAGATGGAAG 495
QY 449 TTGCTATTATTGTCCTTATGATGCGAAGACATTTGAGATTAATCTCCACATT 508
DB 496 TGGCAGTTCTCATCTCTTCCGTAATCGCATGAAACATCTCCAAATTTTCTTACATC 555
QY 509 TGCACCTCGTGTGCGCGCAACAAATTTGACATGATTTTCAATTGAGCAAGTGG 568
DB 556 TTAATTCGAATGCTCCAGAAAGACGCGCTGAAATTTGCTTTATGTCATTGAACACTG 615
QY 569 CGAATCAGACGTTTAATCGCGGAAACTATGAACTTGATGACGATGATCAGCC 628
DB 616 GCACACAACTTTTAACGCTGCGATGCTTTTCATGTGGCTTCAAGAGGCCATGAAG 675
QY 629 TCTACCCATGCGAGCTTCATCTTTCATGATGTCGATTACTGCCCGAAGATGACGTA 688
DB 676 ACAGTGTCTGGGACGTGTATCTTCCAGATGTGATCATCTGAAATGACCGGA 735
QY 689 ACCTGACACGCTGTCATTAACCAACGTCATATGATGATGATGATGATTAATTCATT 748
DB 736 ACTATTTACGATGTGAGAAATGCCAGCTGATTTGCTCAAAAGCTGATTAATATCATGT 795
QY 749 ATAACTTCATATTCGCGCATCTTCCGCGCATCATGCTCACTAACAAGATCATCTGA 808
DB 796 ATATCTTCATATAAAGAAATTTTGGTGGTGAAGTGGGCTGACAGTGAACAATTTA 855
QY 809 AGAAATCATGATTTTGAATGATTTTGGGCTTGGGCGAGAGGACGACGATTTGG 868
DB 856 GAAAGTCATGATGTTTCTTAATGCTTCTGCGGATGCGGAGAGAGATGATGACCTTT 915
QY 869 CGACCGAATCATGATGCTGACGCTGAAGTTTCAAGATATCCGACACAATTTGACGAT 928
DB 916 GGAACGAGATTCATGCTGATGATTAATGTAACGACCAAGAGGAGACTTATGGAATAT 975
QY 929 ATAAATGATTAAGCACTGACGAGAGCGACGATTCAGTTAATATGCTCCGCTACAAA 988

DB 976 ACAAGTCATTTCTCTCA-TCACCATGAGGTGAAGTCCAGTT--TTTAGACGGTATTAAT 1032
QY 989 TAAATGGCCAAACGAAGCCCGCATGACACCTGAGCGCTTAAGCATCTGAAGTATTAAC 1048
DB 1033 TACTAAGGTATTCCAAAGAGCGTCAGTACATGATGAGCTGAACAATTAATATATAGGC 1092
QY 1049 TCGTAATCTGG 1060
DB 1093 CAAAAATCTGG 1104

RESULT 10

ADQ89159
ID ADQ89159 standard; cDNA; 3931 BP.

AC ADQ89159;

DT 21-OCT-2004 (first entry)

DE Human urological disorder related protein 21165 encoding cDNA SEQ:111.

KW urological disorder; uropathic; cytostatic; urinary incontinence;

XX benign prostatic hyperplasia; human; gene; ss.

OS Homo sapiens.

FN Key Location/Qualifiers

FT CDS /tag= a

FT /product= "urological disorder related protein 21165"

XX MO2004065576-A2.

XX 05-AUG-2004.

PF 14-JAN-2004; 2004WO-US000750.

PR 15-JAN-2003; 2003US-0440318P.

PR 04-FEB-2003; 2003US-0444783P.

PR 27-MAR-2003; 2003US-0457901P.

PR 08-MAY-2003; 2003US-0468775P.

PR 19-MAY-2003; 2003US-0471614P.

PR 16-JUN-2003; 2003US-0478742P.

PR 18-JUL-2003; 2003US-0488529P.

PR 30-UTL-2003; 2003US-0491156P.

PR 02-SEP-2003; 2003US-0495944P.

PR 26-SEP-2003; 2003US-0506332P.

XX (MILL-) MILLENNIUM PHARM INC.

PI Karichevi V, Silos-Santiago I, Eliassof SD;

XX WPI; 2004-562167/54.

DR P-PSDB; ADQ89160.

PT Use of polypeptides related to urological disorders, e.g. 44390, 54181,

XX 211 or for identifying a compound capable of treating a urological

PT disorder or identifying and treating a subject having a urological

XX disorder.

Claim 1; SEQ ID NO 111; 542bp; English.

CC The present invention describes the use of polypeptides related to
CC urological disorders for identifying a compound capable of treating a
CC urological disorder, identifying a subject having a urological disorder,
CC or treating a subject having a urological disorder. Also described: (1) a
CC method for identifying a compound capable of treating a urological
CC disorder; (2) a method for identifying a subject having a urological
CC disorder; and (3) a method for treating a subject having a urological
CC disorder. The compound has uropathic and cytostatic activities. The
CC polypeptides related to urological disorders are useful for identifying a
CC compound capable of treating a urological disorder, identifying a subject

CC having a urological disorder, or treating a subject having a urological
CC disorder. Disorders include urinary incontinence and benign prostatic
CC hyperplasia. The present sequence encodes a human urological disorder
CC related protein, which is used in the exemplification of the present
CC invention.

XX Sequence 3931 BP; 1204 A; 725 C; 699 G; 1303 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 13; Length 3931;
Best Local Similarity 54.3%; Pred. No. 1e-40;
Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

```
QY 389 ATCCGACACGACGCGGTGACATGGAATGCTAAGATGTGTGCAAGGCATCGTG 448
DB 710 ATTAGATATTGAGCCAGGGGTCATTTGAGGCCAAAAGACTGTAACCCAGATGGAAG 769
QY 449 TTGCTATTATTTGGCCCTATAGATCTGTAGACACATTTGAGAATATGCTCCAAATT 508
DB 770 TGGCAGTTCATCTCTTCCTTCCTGTAATCGCATGAACATCTTCCAAATTTTCTTACATC 829
QY 509 TGCACGTGTTGCTGGCCAAACAATTTGACATATGCAATTTTTCATTTGTGAGCAAGTGG 568
DB 830 TGAATTCATATGCTCCAGACAGACGGCTGGAATTTGCGTTTATGTATGAAACAGACTG 889
QY 569 CGAATCAGACGTTTAATCGCGGAAACTAATGAACGTTGATACGACATCAAGCC 628
DB 890 GCACACAACCTTTTAACCGTGCATGCTTTTCAATGTGGGCTTCAAGAGGCGCATGAAAG 949
QY 629 TCTAACCATGGCAGTGTCTTATCTTATATGATGCTGATTTACCTCCGAAGATGACCTGA 688
DB 950 ACGAGTCTGGGACTGTGTAATCTTCCACATGTGATCATCTACCTGAAAATGACCGGA 1009
QY 689 ACCTGTACAGTGTCCCAATTCACACGATCATATGATGATGAGCATGCAATTAATTCATT 748
DB 1010 ACTATTTACGATGTGGGAATATGACAGCTCATTTTGTGCAAACTGGATTAATACATGT 1069
QY 749 ATAACTTCATATTCGGCATCTTTCGGCGAATCAGTGCATTAACAAAAGATCAGCTGA 808
DB 1070 ATATTTCTTCATATTAAGAAATTTTGTGTGTGTAATGAGGCTGACAGTGAACAATTTTA 1129
QY 809 AGAAATTCATATGATTTTCCAAATGATTTTGGGGTTGGGGCGAGAGACAGCATTTGG 868
DB 1130 GAAAGATCAATATGTTTCTTAATGCTTCTGGGGATGGGAGGAAGATGATGACCTTT 1189
QY 869 CGACGACATCATGATGCTGACATGAAAGTTTCAAGATATCCGACAAATTCACAGAT 928
DB 1190 GGAACAGATTCACATGCTGATGATTAATTAACCAACAGAGGAGCATTTAGAAAAT 1249
QY 929 ATAAATGATTAAGCACTCGACGAGACGAGAAATCCAGTTAATTAATGCCGCTACAAA 988
DB 1250 ACAAGTCAATTCCTCA-TCAACATAGAGGTGAAGTCCAGTT--TTTAGACGGTATTAAT 1306
QY 989 TAATGGCCCAACGAGAGCGCGATGACAGTGCAGGCTTAAGCAATCTGAAGTAAAGC 1048
DB 1307 TACTTAAGGTATTCACAGAGAGTCAATCATGATGAGCTGAACAAATTTAATATAGGC 1366
QY 1049 TCGTAATCTGG 1060
DB 1367 CAAAATACTGG 1378
```

RESULT 11

ADRA4887 standard; DNA; 3931 BP.

AC ADRA4887;

DT 18-NOV-2004 (first entry)

DE Polypeptide 21165 DNA sequence.

XX de; gene; analgesic; gene therapy; pain; painful disorder.

OS Homo sapiens.

XX P1 WO2004071411-A2.

XX PD 26-AUG-2004.

XX PF 30-JUN-2004; 2004MO-US002851.

XX PR 04-FEB-2003; 2003US-0444781P.

XX PR 05-MAR-2003; 2003US-0452291P.

XX PR 13-MAR-2003; 2003US-0454540P.

XX PR 16-JUN-2003; 2003US-0478805P.

XX PR 30-JUL-2003; 2003US-0491048P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX P1 Silvio-Santiago I, Karicheti V, Eliasof SD;

XX DR WPI; 2004-625760/60.

XX DR P-PSDB; ADRA4888.

XX PS Identifying a compound capable of treating pain by combining a compound

XX CC to be tested with e.g., 16386 polypeptide or with a host cell expressing

XX CC the polypeptide and detecting the binding of the test compound to the

XX CC polypeptide.

XX CC Disclosure; SEQ ID NO 5; 233bp; English.

XX PS The invention relates to a method of identifying a compound capable of

XX CC treating pain or painful disorder by combining a compound to be tested

XX CC with an e.g., 16386, 15402, 21165, 1423, 636, or 32620 polypeptide or with

XX CC a host cell expressing the polypeptide and detecting the binding of the

XX CC test compound to the polypeptide to identify a compound that binds to the

XX CC polypeptide. The method is useful in identifying a compound capable of

XX CC treating pain or painful disorder. This sequence corresponds to the DNA

XX CC encoding polypeptide 21165 - a lactosyl ceramide synthase protein.

XX SQ Sequence 3931 BP; 1204 A; 725 C; 699 G; 1303 T; 0 U; 0 Other;

Query Match 13.9%; Score 159.6; DB 13; Length 3931;

Best Local Similarity 54.3%; Pred. No. 1e-40;

Matches 365; Conservative 0; Mismatches 304; Indels 3; Gaps 2;

```
QY 389 ATCCGACACGACGCGGTGACATGGAATGCTAAGATGTGTGCAAGGCATCGTG 448
DB 710 ATTAGATATTGAGCCAGGGGTCATTTGAGGCCAAAAGACTGTAACCCAGATGGAAG 769
QY 449 TTGCTATTATTTGGCCCTATAGATCTGTAGACACATTTGAGAATATGCTCCAAATT 508
DB 770 TGGCAGTTCATCTCTTCCTTCCTGTAATCGCATGAACATCTTCCAAATTTTCTTACATC 829
QY 509 TGCACGTGTTGCTGGCCAAACAATTTGACATATGCAATTTTTCATTTGTGAGCAAGTGG 568
DB 830 TGAATTCATATGCTCCAGACAGACGGCTGGAATTTGCGTTTATGTATGAAACAGACTG 889
QY 569 CGAATCAGACGTTTAATCGCGGAAACTAATGAACGTTGATACGACATCAAGCC 628
DB 890 GCACACAACCTTTTAACCGTGCATGCTTTTCAATGTGGGCTTCAAGAGGCGCATGAAAG 949
QY 629 TCTAACCATGGCAGTGTCTTATCTTATATGATGCTGATTTACCTCCGAAGATGACCTGA 688
DB 950 ACGAGTCTGGGACTGTGTAATCTTCCAGATGATGATCAATCTTCACTGAAAATGACCGGA 1009
QY 689 ACCTGTACAGTGTCCCAATTCACACGATCATATGATGATGAGCATGCAATTAATTCATT 748
DB 1010 ACTATTTACGATGTGGGAATATGCAAGCTCATTTTGTGCAAACTGGATTAATACATGT 1069
QY 749 ATAACTTCATATTCGGCATCTTTCGGCGAATCAGTGCATTAACAAAAGATCAGCTGA 808
DB 1070 ATATTTCTTCATATTAAGAAATTTTGTGTGTGTAATGAGGCTGACAGTGAACAATTTTA 1129
QY 809 AGAAATTCATATGATTTTCCAAATGATTTTGGGGTTGGGGCGAGAGAGACAGCATTTGG 868
```

Db 1130 GAAAGATCAATGTTTCTTCTTAATGCTTCTTGCGGATGGGAGAGAGATGATGACCTTT 1189
QY 869 CGACGAGAACATCGATGCTGCTGAGCTGAAGTTTCAAGATATCCGACACAAATTTGCACGAT 928
Db 1190 GGAACAGAGTTCACTATGCTGCTGATATATTAATCAACGACCAAGGAGACCTTAGGAAAT 1249
QY 939 ATAAATGATTAAGCACTGACGAGAGCGAGATTCAGTTAATTAATGCCCTACAAAA 988
Db 1250 ACAAGTCATTTCTCTCA-TCAACATAGAGGTGAAGTCACTT--TTTAGACGGTAAAT 1306
QY 989 TAAATGGCCAAACGAGAGCCCGATGACACGTCGCTTAAGCAATCTGAAGTATAGC 1048
Db 1307 TACTAAGGTATTCGAAAGGCGTCAGTACATGATGAGTGAACAATTAATATATAGGC 1366
QY 1049 TCGTAATCTGG 1060
Db 1367 CAAAATCTGG 1378

RESULT 12
AAV73011
ID AAV73011 standard; cDNA; 1853 BP.

XX AAV73011;

DT 15-MAR-1999 (first entry)

DE Human adult brain secreted protein ges1_1 cDNA.

XX Secreted protein; protein factor; human; ges1_1; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 154..975

FT /note= "polynucleotides comprising nucleotides 154-972

XX and 1-341 are also claimed"

XX MO9842741-A2.

PD 01-OCT-1998.

XX 25-MAR-1998; 98WO-US005972.

XX 25-MAR-1997; 97US-00825145.

XX 24-MAR-1998; 98US-00046881.

XX (GENY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;

XX Spaulding V, Agoestino MJ;

XX WPI; 1999-045135/04.

XX P-PSDB; AAW82006.

XX New polynucleotides encoding secreted human proteins - are derived from

XX human foetal brain, adult testes, adult brain, or placenta cDNA

XX libraries, potentially useful as, e.g. immunostimulants or

XX antiinflammatory agents.

XX Claim 35a; Page 85-86; 113pp; English.

XX This cDNA sequence codes for novel human secreted protein ges1_1 (see

XX AAW82006). The cDNA was isolated from a human adult brain cDNA library

XX using methods which are selective for cDNAs encoding secreted proteins,

XX or was identified as encoding a secreted or transmembrane protein on the

XX basis of computer analysis of the amino acid sequence of the encoded

XX protein. The nucleotide sequence shows homology to some known database

XX sequences. ges1_1 is a full-length clone and can be obtained from the

XX deposit clone ATCC 98371 as a 1850 bp EcoRI/KcoI fragment. The invention

XX provides polynucleotides (see AAV7300-12), derived from human foetal

XX brain, adult testis, adult brain and adult placenta cDNA libraries, that

CC encode novel secreted proteins (see AAW81997-07). These can be isolated
CC from composite clone ATCC 98371 using specific probes (see AAV73013-22).
CC The polynucleotides and proteins are predicted to have biological
CC activities which would make them suitable for treating, preventing or
CC ameliorating medical conditions in humans and animals (no supporting data
CC are given). Suggested activities include nutritional, immune stimulating
CC (e.g. as vaccines) or suppressing, haematopoietic, tissue
CC growth, activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour invasion
CC suppressor, and tumour inhibition activities. The polynucleotides are
CC also stated to be useful for gene therapy

XX Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;

Query Match 13.6%; Score 157.2; DB 2; Length 1853;
Best Local Similarity 54.4%; Pred. No. 4,2e-40;
Matches 360; Conservative 0; Mismatches 298; Indels 4; Gaps 2;

QY 421 CCTAAGATTTGTTGCAAGGATGCTGCTATTTATTTGTCCTTATAGAGATGCTGAA 480

Db 499 CCGTGAAGTTGTGAGACCCCGCTCCGAAACAGCATATTTGCTCATGTCGCCGGAG 558

QY 481 GCACTTTGAGATATAGCTCCCAATTTGCACTGCTGCTGCCAAACAAATTTGAC 540

Db 559 CACCACTGCGCTGCTGCTCTTACCACTGCACTCCCTTCTGACGCGCAGCAGCTTGT 618

QY 541 TATGCAATTTTCAATGTTGAGCAAGTGGCGAATGCAAGCTTAAATCGGGGAACTAATG 600

Db 619 TATGCAATTTTCAATGTTGAGCAAGTGGCGAATGCAAGCTTAAATCGGGGAACTAATG 678

QY 601 AACGTTGATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

Db 679 AACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 738

QY 661 GTGATTTACTGCTCCGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 717

Db 729 GTGATTTACTGCTCCGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 798

QY 718 CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 777

Db 799 CATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858

QY 778 GGAATCAGTGCATCAACAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 837

Db 859 AGAGTCTGAGCACTTACCTGACCAAGATGATGATGATGATGATGATGATGATGATGAT 918

QY 838 TGGGTTGGGGGAG 897

Db 919 TGGGTTGGGGGAG 977

QY 898 GTTTCAGATATCCGACCAATTTGCAAGATGATGATGATGATGATGATGATGATGATGAT 957

Db 978 ATCTCTGAGCCGCCCACTCTGTAGACATTAATGATGATGATGATGATGATGATGATGAT 1037

QY 958 ACGAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1017

Db 1038 GCGAATGAGAAATATCCCAAGATTTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1097

QY 1018 CGTGAGCGCTTAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077

Db 1098 CAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1157

QY 1078 AC 1079

Db 1158 AC 1159

RESULT 13
ABQ92036
ID ABQ92036 standard; cDNA; 1853 BP.

XX ABQ92036;

XX ABQ92036;

DT 04-OCT-2002 (first entry)
 XX Human polynucleotide SEQ ID NO 33.
 DE
 XX Human; cytostatic; antirheumatic; antiarthritic; vulnary; analgesic;
 KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KW anticancer; fungicide; antidiabetic; antiaesthetic; antiallergic;
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KW stem cell; growth factor; nervous system disease; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 PN US2002065394-A1.
 XX
 PD 30-MAY-2002.
 XX
 XX 22-DEC-2000; 2000US-00745763.
 XX
 XX 18-MAR-1998; 98US-00040963.
 XX
 PA (JACO/) JACOBS K.
 PA (MCCO/) MCCOY J. M.
 PA (LAVA/) LAVALLIE E. R.
 PA (COLL/) COLLINS-RACIE L. A.
 PA (EVAN/) EVANS C.
 PA (MERB/) MERBERG D.
 PA (TREA/) TREACY M.
 PA (SPAU/) SPAULDING V.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Spaulding V;
 DR P-PSDB; ABP61819.
 XX
 PT Novel secreted or transmembrane protein and polynucleotide encoding the
 PT protein, useful for diagnosis and treatment of neurological disorders,
 PT cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.
 XX
 PS Claim 11; Page 144-145; 284pp; English.
 XX
 CC The invention relates to human secreted or transmembrane protein (I),
 CC their fragments and is encoded by specific complementary deoxyribonucleic
 CC acid (cDNA) inserts (II), where the protein is substantially free from
 CC other mammalian proteins. (I) are useful for preventing, treating or
 CC ameliorating a medical condition, especially immunological treatment or
 CC prevention of tumours. (I) exhibits activity relating to angiogenesis,
 CC cytokine, cell proliferation, cell differentiation, antiinflammatory,
 CC stem cell growth factor activity and activin or inhibin-related
 CC activities. (I) can be used to manipulate stem cells in culture to give
 CC rise to neuroepithelial cells that can be used to augment or replace
 CC cells damaged by illness, autoimmune disease, accidental damage or
 CC genetic disorders. (I) induces the proliferation of neural cells and
 CC regeneration of nerve and brain tissue and is useful for the treatment of
 CC central and peripheral nervous system diseases and neuropathies, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, Amyotrophic
 CC lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis and is useful for treating myeloid
 CC or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 CC and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 CC tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 CC for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 CC periodontal disease. (I) is also useful for gut protection or
 CC regeneration and treatment of lung or liver fibrosis, reperfusion injury
 CC in various tissues, various immune deficiencies and disorders including
 CC severe combined immunodeficiency (SCID), bacterial or fungal infections,
 CC autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 CC diabetes mellitus, myasthenia gravis, allergic reactions and conditions,

CC such as asthma or other respiratory problems. (II) is useful to express
 CC recombinant protein, as markers for tissues in which the corresponding
 CC protein is preferentially expressed and in gene therapy. The present
 CC sequence is that of a polynucleotide of the invention

XX Sequence 1853 BP; 372 A; 540 C; 483 G; 450 T; 0 U; 8 Other;

Query Match 13.6%; Score 157.2; DB 6; Length 1853;
 Best Local Similarity 54.4%; Pred. No. 4.2e-40;
 Matches 360; Conservative 0; Mismatches 298; Indels 4; Gaps 2;

421 CCTAAGATTGTTGTTCAAGGCGATGTTGCTATTATTTGTCCTTATAGATCGTGA 480
 DB 499 CCGGAGGTTGAGGCCCGCTCCGAAACAGCATATGTCCTATGTCGCCGGAG 558
 QY 481 GCACATTGAGATATATGCTCCCAATTTGACCTGCTGCCCAACAAATTTGAC 540
 DB 559 CACCACTTGGCGCTGCTGCTTCAACCTGACACCTTTTTCAGAGCCAGCTTGT 618
 QY 541 TATGCAATTTTCAATTTGAGAGCAAGTGGCAATTCAGACGTTTAAATGCGGAAACTAATG 600
 DB 619 TATGGCATCATGTATCATCCAGAGCTGGAATGGAACATTTTAAACAGGGCAAACTGTTG 678
 QY 601 AACGTTGATPACAGCTAGCATCAGCCTCTACCAATGGCAGTCTTCAATGAT 660
 DB 679 AACGTTGGGGTGGAGAGGCCCTGCTGATGGAAGATGGAATGCTGTTTGGACGAT 738
 QY 661 GTGATTTTACGCGGAGAGTACCGTAACTGTACACGTGT---CCAATTCACACG 717
 DB 739 GTGGACCTCTTGGCAAAATGACCAAACTGTATGTGTGACCCCGGGGACCCCG 798
 QY 718 CATATGATGTAGCGATCGATCAATTAATTAATTTTCAATTTTGGCGATCTTGGC 777
 DB 799 CATGTTGCCGTTGCTATGAAACAAATTGGATACAGCTCCGTAACCCCAAGTACTTGG 858
 QY 778 GGAATAGTGCATCAACAAAGTCACTGGAAGAAATCAATGATTTTGAATGATTT 837
 DB 859 AGAGTCTACGACATTTACTCTGTGACGATACCTGAAGATGAATGCTTCCCAATGATAC 918
 QY 838 TGGGGTTGGGGCGGAGAGACGACGATTTGGCGACAGAAATCATGATGCTGACTGAA 897
 DB 919 TGGGGCTGGGGTGTGAGATGACGACTT-GCTACAGGGTGGCGCTGGTGGATGAA 977
 QY 898 GTTTCAGATATTCGACACAAATTTGACGATATTAATTAATTAAGACTGACGAGCG 957
 DB 978 ATCTCTGGGCCCCCATCATCTGTAGGACACTATAAGATGATGAAGCAGAGATTAAG 1037
 QY 958 ACGAATTCAGTTAATTAATGCCCTACAAATTAATGGCCAAACGAAGCCGATGACA 1017
 DB 1038 GGCMAATGAGAAATATCCCAAGATTGACCTCTGTCGTAACCGAATTCCTGACG 1097
 QY 1018 CGTAGCGGCTTAGCATCTGAATTAAGCTGTAAATCTGGAATTTGAAGCTCTCTAC 1077
 DB 1098 CAAGATGGGATGATCACTGATCAATACAGTCTGCTGAGAGCTGGGCTCTTTAT 1157
 QY 1078 AC 1079
 DB 1158 AC 1159

RESULT 14
 ID ADO31358 standard; DNA; 1035 BP.
 XX

AC ADO31358;

DT 07-OCT-2004 (first entry)

DE Human transferase coding sequence, SEQ ID 1.

XX Human; transferase; enzyme;

KW N-acetylglucosamine-6-O sulfate-specific beta 1;
 KW 4-galactosyl transferase; beta 1; 4-galactosylation; gene; ds.


```
PR 12-JUN-1998; 98US-0089105P.
PR 16-JUN-1998; 98US-0089440P.
PR 16-JUN-1998; 98US-0089512P.
PR 16-JUN-1998; 98US-0089514P.
PR 17-JUN-1998; 98US-0089532P.
PR 17-JUN-1998; 98US-0089538P.
PR 17-JUN-1998; 98US-0089588P.
PR 17-JUN-1998; 98US-0089598P.
PR 17-JUN-1998; 98US-0089600P.
PR 17-JUN-1998; 98US-0089653P.
PR 18-JUN-1998; 98US-0089601P.
PR 18-JUN-1998; 98US-0089607P.
PR 18-JUN-1998; 98US-0089908P.
PR 19-JUN-1998; 98US-0089947P.
PR 19-JUN-1998; 98US-0089948P.
PR 19-JUN-1998; 98US-0089952P.
PR 22-JUN-1998; 98US-0090246P.
PR 22-JUN-1998; 98US-0090254P.
PR 22-JUN-1998; 98US-0090349P.
PR 23-JUN-1998; 98US-0090355P.
PR 24-JUN-1998; 98US-0090429P.
PR 24-JUN-1998; 98US-0090431P.
PR 24-JUN-1998; 98US-0090435P.
PR 24-JUN-1998; 98US-0090444P.
PR 24-JUN-1998; 98US-0090445P.
PR 24-JUN-1998; 98US-0090461P.
PR 24-JUN-1998; 98US-0090472P.
PR 24-JUN-1998; 98US-0090535P.
PR 24-JUN-1998; 98US-0090538P.
PR 24-JUN-1998; 98US-0090540P.
PR 25-JUN-1998; 98US-0090557P.
PR 25-JUN-1998; 98US-0090678P.
PR 25-JUN-1998; 98US-0090688P.
PR 25-JUN-1998; 98US-0090690P.
PR 25-JUN-1998; 98US-0090691P.
PR 25-JUN-1998; 98US-0090694P.
PR 25-JUN-1998; 98US-0090695P.
PR 25-JUN-1998; 98US-0090696P.
PR 26-JUN-1998; 98US-0090862P.
PR 26-JUN-1998; 98US-0090863P.
PR 01-JUL-1998; 98US-0091358P.
PR 01-JUL-1998; 98US-0091360P.
PR 02-JUL-1998; 98US-0091478P.
PR 02-JUL-1998; 98US-0091486P.
PR 02-JUL-1998; 98US-0091519P.
PR 02-JUL-1998; 98US-0091544P.
PR 02-JUL-1998; 98US-0091626P.
PR 02-JUL-1998; 98US-0091628P.
PR 02-JUL-1998; 98US-0091633P.
PR 02-JUL-1998; 98US-0091646P.
PR 02-JUL-1998; 98US-0091673P.
PR 07-JUL-1998; 98US-0091978P.
PR 07-JUL-1998; 98US-0091982P.
PR 07-JUL-1998; 98US-0092182P.
PR 10-JUL-1998; 98US-0092472P.
PR 10-JUL-1998; 98US-0093339P.
PR 30-JUL-1998; 98US-0094651P.
PR 04-AUG-1998; 98US-0095282P.
PR 04-AUG-1998; 98US-0095285P.
PR 04-AUG-1998; 98US-0095301P.
PR 04-AUG-1998; 98US-0095302P.
PR 04-AUG-1998; 98US-0095318P.
PR 04-AUG-1998; 98US-0095321P.
PR 04-AUG-1998; 98US-0095325P.
PR 10-AUG-1998; 98US-0095916P.
PR 10-AUG-1998; 98US-0095929P.
PR 10-AUG-1998; 98US-0096012P.
PR 11-AUG-1998; 98US-0096143P.
PR 11-AUG-1998; 98US-0096146P.
PR 12-AUG-1998; 98US-0096329P.
PR 17-AUG-1998; 98US-0096757P.

PR 17-AUG-1998; 98US-0096766P.
PR 17-AUG-1998; 98US-0096768P.
PR 17-AUG-1998; 98US-0096773P.
PR 17-AUG-1998; 98US-0096791P.
PR 17-AUG-1998; 98US-0096867P.
PR 17-AUG-1998; 98US-0096891P.
PR 17-AUG-1998; 98US-0096894P.
PR 17-AUG-1998; 98US-0096895P.
PR 17-AUG-1998; 98US-0096897P.
PR 18-AUG-1998; 98US-0096949P.
PR 18-AUG-1998; 98US-0096950P.
PR 18-AUG-1998; 98US-0096959P.
PR 18-AUG-1998; 98US-0096960P.
PR 18-AUG-1998; 98US-0097022P.
PR 19-AUG-1998; 98US-0097141P.
PR 20-AUG-1998; 98US-0097218P.
PR 24-AUG-1998; 98US-0097661P.
PR 26-AUG-1998; 98US-0097951P.
PR 26-AUG-1998; 98US-0097952P.
PR 26-AUG-1998; 98US-0097954P.
PR 26-AUG-1998; 98US-0097955P.
PR 26-AUG-1998; 98US-0097971P.
PR 26-AUG-1998; 98US-0097974P.
PR 26-AUG-1998; 98US-0097978P.
PR 26-AUG-1998; 98US-0097979P.
PR 26-AUG-1998; 98US-0097986P.
PR 26-AUG-1998; 98US-0098014P.
PR 31-AUG-1998; 98US-0098525P.
PR 16-SEP-1998; 98US-0100634P.
PR 12-JAN-1999; 99US-0115565P.

XX (GETH ) GENENTECH INC.
XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;
PI Wood WI, Yuan J;
XX WPI; 2000-072883/06.
DR P-PSDB; AAY66696.
XX Membrane-bound proteins and related nucleotide sequences.
PT
XX Claim 2; Fig 160; 822pp; English.
XX
XX The invention provides membrane-bound PRO polypeptides and
CC polynucleotides encoding them. The PRO sequences of the invention were
CC identified based on extracellular domain homology screening. The PRO
CC sequences have homology with proteins including LDL receptors, TIE
CC ligands and various enzymes. The membrane-bound proteins and receptor
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor
CC immunoadhesins, for instance, can be used as therapeutic agents to block
CC receptor-ligand interactions. The membrane-bound proteins can also be
CC employed for screening of potential peptide or small molecule inhibitors
CC of the relevant receptor/ligand interaction. The PRO encoding sequences
CC are useful as hybridization probes, in chromosome and gene mapping and in
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences will
CC also be useful for the preparation of PRO polypeptides, especially by
CC recombinant techniques
XX
SQ Sequence 1964 BP; 592 A; 365 C; 474 G; 533 T; 0 U; 0 Other;

Query Match 13.5%; Score 155.4; DB 3; Length 1964;
Best Local Similarity 51.8%; Pred. No. 1.7e-39;
Matches 351; Conservative 0; Mismatches 326; Indels 0; Gaps 0;

QY 421 CCTAAGATTGTTGTCGAAGGCATCGTGTGCTATTAATTTGCGCCCTAATAGATCGTGA 480
DB 461 CCTCAGGAATGTAACTTTACAGAGGTCGCATCTCTTCCACCGGAACAGAG 520
QY 481 GCACATTGAGTAATTAAGTCCACATTTGCACTCGTGTGCTCCCAACACATTTGAC 540
DB 521 AAACACCTGATGTACTCTGCGAACAATCTGCATCCCTTCTCGAGAGGACGACGTGAT 580
QY 541 TATGCAATTTTCATTGTGAGCAAGTGGCGAATCAGCGTTTAATCGGGAAACTAATG 600
```

Db 581 TATGGCATCTACGTCATCCACGAGGCTGAAGGTAAAGTTAATGAGCCAAACTCTTG 640
Qy 601 AACGTTGATAGAGCTAGCATCACGCCCTCTACCCATGGCAGTGTCTCATCTTTATGAT 660
Db 641 AATGGGGCTATCTAGAACCCCTCAAGAGAAATTTGGAGCTCTTTATATTCACGAT 700
Qy 661 GTGATTTTACTGCCCCGAAATGACCGTAACTGTACACGTGTCCAAATTCACCAACGTCAT 720
Db 701 GTGGACCTGATCCGAGAAATGACTTTAACCTTACAAAGTGAAGAGCATCCCAAGCAT 760
Qy 721 ATGAGTGAAGCATGATTAATTAATTCATTTAACTTCATATTCGCGATCTTTCGCGGA 780
Db 761 CTGGTGTGGCAGGAACAGCACTGGGTACAGGTACGTTACAGTGATATTTTGGGGGT 820
Qy 781 ATCAGTGCCTAACAAAGATCACTGAAGAAATCAATGATTTTCAATGATTTTGG 840
Db 821 GTTACTGCCCTTAGCGAGAGAGAGATTTTCAAGGTGAATGATTCCTAACAACTACTGG 880
Qy 841 GGTGGGGCGAGAGAGAGAGATTTGGGACGAAACATCGATGGCTGACTGAAGTT 900
Db 881 GATGGGGAGGCGAGAGACATGACTCAGACTCAGGGTTGAGCTCCAAAGAAATGAATTT 940
Qy 901 TCAAGATATCCGACACAAATTCACGATATTAATGATTAGCACTCGAAGGCGACG 960
Db 941 TCCCGGCCCTGCTGAAGTGGGTAAATATACATGCTCTTCACACTAGAGCAAAAGGC 1000
Qy 961 AATCCAGTTAATAATGCCGCTACAAATATATGGCCAAACGAGGCCGATGACACGT 1020
Db 1001 AATGAGGTGAACGACGAACGATGAAGCTTTACACCAAGTGCACGAGTCTGAGAAC 1060
Qy 1021 GACGGCTTAAGCAATCTGAAGTATAGCTGTAATCTGAAATTAAGCCTCTACACT 1080
Db 1061 GATGGTTGAGTGTGTTCTTATTAATAGTATCTGTGAAACACAACTCTTATATATC 1120
Qy 1081 CGAGCCGTGTCGATTT 1097
Db 1121 AACATCAGAGTGATTT 1137

Search completed: January 8, 2006, 19:19:43
Job time : 783 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 18:40:30 ; Search time 5190 Seconds
(without alignments)
10385.110 Million cell updates/sec

Title: US-10-661-430-2
Perfect score: 1152
Sequence: 1 atgcttcctgcatttgcgc.....acttccaacgtttttag 1152

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	672.4	58.4	860	7	CV125668
2	643	55.8	866	7	CV125689
3	178.2	15.5	687	5	BM054310
4	177.8	15.4	679	3	BM061791
5	177.2	15.4	622	5	BM347970
6	167.6	14.5	796	6	CX757659
7	167.6	14.5	879	6	CD514077
8	167.6	14.5	955	5	BQ957788
9	167.6	14.5	1196	11	DQ035706
10	167.6	14.5	1590	4	CR614082
11	167.6	14.5	1860	4	CR612341
12	167.6	14.5	1870	4	CR607487
13	167.6	14.5	1876	4	CR582876
14	167.6	14.5	1876	4	CR594939
15	165	14.3	673	4	AJ740626
16	164.4	14.3	1979	4	CR925982
17	163.6	14.2	402	1	AI987157
18	163.6	14.2	672	6	CD081993
19	163.4	14.2	715	7	CN555132
20	163.2	14.2	612	7	CK546970
21	162.4	14.1	667	5	BM052849
22	161.2	14.0	893	8	CX362564

23	161	14.0	661	3	BU617372	BU617372
24	160	13.9	588	5	BU029426	BU029426
25	160	13.9	654	3	BM437339	BM437339
26	157.6	13.7	970	1	AL558425	AL558425
27	157.2	13.6	844	1	AL866706	AL866706
28	155.4	13.5	1035	10	AY410571	AY410571
29	155.4	13.5	2303	4	CR749555	CR749555
30	155	13.5	561	5	BM044315	BM044315
31	155	13.5	803	8	CX744001	CX744001
32	154.4	13.4	822	8	DN948717	DN948717
33	153.4	13.3	737	8	CX744145	CX744145
34	152.8	13.3	774	7	CK026161	CK026161
35	152.4	13.2	1149	10	AY412728	AY412728
36	152.2	13.2	907	5	BM489994	BM489994
37	151.6	13.2	690	7	CN280919	CN280919
38	151.4	13.1	791	8	CX780748	CX780748
39	151.2	13.1	831	5	BQ918131	BQ918131
40	150	13.0	759	8	CX457513	CX457513
41	150	13.0	1052	10	AY417315	AY417315
42	150	13.0	1368	4	CN50GMSO	CR725570
43	149	12.9	1009	10	AY410572	AY410572
44	148.8	12.9	1738	4	BC004095	BC004095
45	148.8	12.9	2369	4	BC029303	BC029303

ALIGNMENTS

RESULT 1
CV125668 860 bp mRNA linear EST 31-AUG-2004
LOCUS OSTR30043C08 PDONR201 Entry vector Caenorhabditis elegans cDNA,
DEFINITION mRNA sequence.
ACCESSION CV125668.1 GI:51716887
VERSION EST.
KEYWORDS Caenorhabditis elegans
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE 1 (bases 1 to 860)
Lamesch, P., Milstein, S., Hao, T., Rosenberg, J., Li, N., Sequerra, R., Bosak, S., Doucette-Stamm, L., Vandenhaute, J., Hill, D.E. and Vidal, M.
C. elegans ORFeome Version 3.1: Increasing the coverage of ORFeome resources with improved gene predictions
Genome Res. (ORFeome issue) (2004) In press
JOURNAL Contact: Philippe Lamesch and Tong Hao
COMMENT Marc Vidal Lab
DFCI

44, Binney Street, Boston, MA 02115, USA
Tel: 6176323910
Fax: 6176325739
Email: philippe_lamesch@dfci.harvard.edu
PCR Primers
FORWARD: TGGCTTTCGTCATTGCGCAGT
BACKWARD: TAAACACGCTGGAAGTCCCT.
location/Qualifiers
1..860
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="male and hermaphrodite"
/tissue_type="all tissues"
/cell_type="all cells"
/dev_stage="all stages"
/clone_lib="PDONR201 Entry vector"

ORIGIN

Query Match 58.4%; Score 672.4; DB 7; Length 860;
Best Local Similarity 97.9%; Pred. No. 1.1e-202;
Matches 703; Conservative 0; Mismatches 11; Indels 4; Gaps 2;

DEFINITION BM054310 Nori Satoh unpublished cDNA library, blood cells Ciona intestinalis cDNA clone c1db080a21 5', mRNA sequence.

ACCESSION BM054310

VERSION BM054310.1 GI:24155006

KEYWORDS EST

SOURCE Ciona intestinalis

ORGANISM Ciona intestinalis

REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Ciona.

AUTHORS Satou, Y., Satake, M., Azumi, K., Nonaka, M., Shin-i, T., Kohara, Y. and Satoh, N.

TITLE Expressed genes in Ciona intestinalis (2002)

JOURNAL Unpublished (2002)

COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.

FEATURES

source location/Qualifiers

1..687

/organism="Ciona intestinalis"

/mol_type="mRNA"

/db_xref="taxon:7719"

/clone="c1db080a21"

/tissue_type="blood cells"

/clone_11b="Nori Satoh unpublished cDNA library, blood cells"

ORIGIN

Query Match 15.5%; Score 178.2; DB 5; Length 687;
Best Local Similarity 55.1%; Pred. No. 7.2e-45;
Matches 348; Conservative 0; Mismatches 284; Indels 0; Gaps 0;

448 GTTGCTATTATTGTCCTTATAGATCGTGAAGACATTTGAGATATGCTCCACAT 507
|||||
30 GTTGCAATGTTATCTTATCGGGACAGAGAAGACATTTGCGGATTTCTCGAGTAC 89
|||||
508 TTGCACTGCTGCTGCCCAACACATTTGACTATGCAATTTTCACTTGGAGCAAGTG 567
|||||
90 ATGCAATCCCAACCCCTCAACAGTCAACACTGATTAAGCCATCTATGTTATACAGGCT 149
|||||
568 GCGAATCAGAGCTTAAATGCGGGAACTAATGAACGTTGATACAGAGTACATCAGCC 627
150 GGTACTGCTAAATTTATGTCGCAAGCTGATGAATGTTGTTATGCTGATCAATCAAA 209
|||||
628 CTCTACCCATGCGAGTCTTCACTTTTCATGATGTCGATTTTCTGCCGAGATGACGCT 687
210 GATCATGATTTTCATGTTTTCCTTTCATGATGTCGATTTGCTGCTTGAAGATGACAAG 269
|||||
688 AACCTGTACAGCTGTCCCAATTCACACGCTCATATGATGTAAGCATCGATTAATTCAT 747
270 AGTATATCTCTTGGCCCCCAAGTCTCGCCACCTGTCAGCTGGGGTGAACAAGTTTAT 329
|||||
748 TATAAATCTCCATTTTCGCGGATCTTCGCGGGAATCAGTCACTAAACAAATACCTCG 807
330 TACCAACTGCCCCCTACTCTGCATATTTGATGCGTACTGAGCTGCAAGAGAACATTC 389
|||||
808 AAGAAATCAATGATATTTTCGATGATTTTTCGGGTTGGGGCGAGAGACACGATTTG 867
390 CAAAAGTGAAGGTTTACTCTAATTCATTTCTGGGGTTGGGGTGAAGAGACACATG 449
|||||
868 GCGAGCAACATCGATGCTGCACTGAAGTTTCAAGATATCGACACAAATTCGACGA 927
450 TTTAATCCGTCAGATTTTCTGCGCATGATATTTATTCCTTACCTATGATATCTCAAG 509
|||||
928 TATAAATGATTTAGACATCGACGAGAGAGCAATTCAGTTAATTAATCCGCTACAA 987
510 TATAAGATATTTACTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 569
988 ATTAATGGGCAACGAGCGCGATGACAGTGAAGCGCTTAAGCAATCTGAAGTAAAG 1047

Db 570 CAATATCCAGGACTTAAGACACGATGGCAACGATGCTTAACACATTAATACNAG 629
|||||

Qy 1048 CTCGTAAATCTGATTTGAAGCTCTCTACAC 1079
|||||

Db 630 GTCTATCAAGACAGAGAGACAGCTTTATAC 661
|||||

RESULT 4

LOCUS B061791 679 bp mRNA linear EST 29-SEP-2003

DEFINITION B061791 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone X1074b05 5', mRNA sequence.

ACCESSION B061791

VERSION B061791.1 GI:17423856

KEYWORDS EST

SOURCE Xenopus laevis (African clawed frog)

ORGANISM Xenopus laevis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus; Xenopus.

AUTHORS Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.

TITLE Expressed genes in X. laevis embryo

JOURNAL Unpublished (2001)

COMMENT Contact: Tadasi Shin-i
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@gens.nig.ac.jp
The information of this clone is available through the following URL.
http://xenopus.nibb.ac.jp.

FEATURES

source location/Qualifiers

1..679

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="X1074b05"

/tissue_type="whole embryo"

/dev_stage="stage 25"

/clone_11b="NIBB Mochii normalized Xenopus tailbud library"

ORIGIN

Query Match 15.4%; Score 177.8; DB 3; Length 679;
Best Local Similarity 53.8%; Pred. No. 9.6e-45;
Matches 365; Conservative 0; Mismatches 313; Indels 0; Gaps 0;

418 ATGCTAAGATGTTGTTGCAAGCATGCTGTTCTATTATTTGTCCTTAAGATGCT 477
|||||
2 AGGCCCGACGATTTGATGATGCCAGCAGAGAGAGTGGCATTAATCCATTCGATGCA 61
|||||
478 GAAGCACTTTGAAGATTAATGTCACATTTGCACTGCTGCTGCCAACAACAATTG 537
62 GAAGCTCATCTCAAGACATGCTGTAATCAAGCACCCTTCTGACGCGGACAGCA 121
|||||
538 GACTATGCAATTTTCATTTGAGAGCAAGTGGCAATCAGAGCTTAATTCGCGGAACTA 597
122 GATTATGTTGATATATGTTGTTGGAACAGATGGAAGACAGATTTTCAACGACCAAGCTG 181
598 ATGAAGTTTGATACAGACATGATCAGCCTCTTACCAATGAGAGTCTTATCTTAT 657
182 ATGATGTTGTTATGCGACAGCATCAAAAGACTATGATTTATGCTGTTTATATTCAGT 241
|||||
658 GATGCGATTTACTGCCGAGAGATGACGTAACGTTACGTCAGTGTCAATTCACAGCT 717
242 GATGCGACATTTATCTTATGATGAGAGAAATCTTACAGGCTCTTCAAAATCCAGCA 301
718 CATATAGTGTAGAGATGATTAATTCATTAATTAATTCATTTCCGCGATCTTCGCG 777

```
Db      302 CATATGGCAAAATTCCTGTGACAAAGTTAACTTTAAATTAACCTTACATGACATCTTTGGA 361
Qy      778 GGAATGACGTGACCTAAACAAAAGATCACTGAGAGAAAATCAATGATTTTGAATGATTTT 837
Db      362 GGAAGTGTGGCATTTCAAAAAGAACCTTTATCAAAAGTGAATGGAATTTCTCAATGATTTTC 421
Qy      838 TGGGGGTGGGGGAGAGAGAGCGATTTTGGCGAGCAGAAACATCGATNGCTGGACTGAAA 897
Db      422 TGGGGGTGGGGAGAGAGATGATGAACTGTTCCAAAGAGCTGTGGCTGGAGCTGAAA 481
Qy      898 GTTTCAGATATCCGACACAAATTTGACATATTAATTAATTAAGATCTGACGAGAGCG 957
Db      482 GTTGAACGCTGCAGACCAACAAATTTGCCAATCCAGATGATTTCTCATAAAAGATTCCT 541
Qy      958 AGCAATCCGATTATTAATGCGCTGACAAATATATGGCCCAAGAGAGCGCGATGACA 1017
Db      542 GGTAAATGAAGAATGAAGAGCTTTCTCTAATTAAGAAAGCCCTATTAACCTTCGCAT 601
Qy      1018 CGTGAAGGCTTAGCAATCTGAAGTAAAGCTGTAATCTGGAATTTGAAGCCTCTCTAC 1077
Db      602 CAGAGATGAGACTGATTCCTGGAATTAACCATATATAGTAACTAGTCAATAACTTTAT 661
Qy      1078 ACTGAGCCGCTGCTGCAT 1095
Db      662 ACAAAAGTCACAGTGAT 679

RESULT 5
LOCUS      BM347970 622 bp mRNA linear EST 27-MAY-2004
DEFINITION BM347970 Yutaka Satou unpublished cDNA library, embryo whole animal
            Ciona intestinalis cDNA clone ciem838b10 5', mRNA sequence.
ACCESSION  BM347970
VERSION     BM347970.1 GI:47759771
KEYWORDS    EST.
SOURCE      Ciona intestinalis
ORGANISM    Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Clonidae; Ciona.

REFERENCE  1 (bases 1 to 622)
AUTHORS   Satou,Y., Shin-I,T., Kohara,Y. and Satoh,N.
TITLE      Expressed genes in Ciona intestinalis (2004)
JOURNAL    Unpublished (2004)
COMMENT    Contact: Yutaka Satou
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4095
            Fax: 81-75-705-1113
            Email: yutaka@ascidian.zool.kyoto-u.ac.jp.

FEATURES
            source
            1..622
               .organism="Ciona intestinalis"
               /mol_type="mRNA"
               /db_xref="taxon:7719"
               /clone="ciem838b10"
               /tissue_type="whole animal"
               /dev_stage="embryo"
               /clone_id="Yutaka Satou unpublished cDNA library, embryo
               whole animal"
```

Query Match 15.4%; Score 177.2; DB 5; Length 622;
Best Local Similarity 55.8%; Pred. No. 1.4e-44;
Matches 338; Conservative 0; Mismatches 268; Indels 0; Gaps 0;

Qy 448 GTTGCTATTATTGGCCCTATAGATCGTGAAGACATTTGAGATTAAGCTGCACAAT 507
Db 12 GTTGCCATAGTATTCTTATCGGACAGACAAACATTTGCGGTATTTCTCTGAGTAC 71
Qy 508 TTGCACTCGTTGCTGCGCAACAAATGGACTATGCAATTTTCATTTGGAGCAAGTG 567

```
Db      72 ATGATCCCAACCCCTCAACAGTCACAACTTGATTAAGCCATCTATGTTGTTATGAGCT 131
Qy      568 GGAATGACAGCGTTTAATGCGGGGAAACCTAATGAACGTTGATGACGATGATCAGC 627
Db      132 GGTACTGGTAAATTTAATGATGTCACAGCTGATGAATGTTGGTTATGCTGATGATCAAA 191
Qy      628 CTCTACCCATGGCAGTGTCTTCACTCTTCATGATGATGATTTATGCCCCGAGATGACCGT 687
Db      192 GATCATGATTTTCATAGTTTCTTCTTCCATGATGATGAGACTTGCTGTTGAGATGACAAG 251
Qy      688 AACCTGTACAGGTGTCGAATTAACACAGCTCATATGAGTGTAGCTGATGATTAATTCAT 747
Db      252 AGTATATACTCTTGGCCCCCTCAAGTCTCCGACCTGTACGCTGGGGGTGACAAAGTTTAT 311
Qy      748 TATTAACCTTCATTTTGGCGATCTTGGCGGAATCAAGTACGACCTAACAAAGATCAGCTG 807
Db      312 TACCAAGTCCCTTACTCTGCAATTTGGTGGCGTGAAGTGAAGTGAAGTGAAGTGAAG 371
Qy      808 AAGAAATCAATGATTTTTCGAATGATTTTGGGGTTGGGGCGGAGAGAGACGATTTTG 867
Db      372 CAAAAGTGAAGGTTTACTCTAATTCATTTCTGGGGTTGGGGTGGAGAGACGATATG 431
Qy      868 GCGACGAAACATCGATGCGCTGGAAGTGAAGTTCAGATATCCGACAAATTTGACGA 927
Db      432 TTTAACCGTGTCAAGTTTCTGGCATGAATATTTGTTTACCTATGATATCTCAAG 491
Qy      928 TATTAATGATTAAGCACTGACGAGAGGAGCAAGATCCAGTTAATTAATGCGGCTCAAA 987
Db      492 TATTAATGATTAATCTCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 551
Qy      988 ATTAATGGGCAACGAAAGCGCGATGACACGTCGACGCTTAAGCAATCTGAAGTAAAG 1047
Db      552 CAATACGACGAGCTTAAGACACAGATGCAAGAGATGCTTAACACACTGATTAATCAAG 611
Qy      1048 CTCGTA 1053
Db      612 GTCGTA 617
```

RESULT 6
LOCUS CX757659 796 bp mRNA linear EST 24-JAN-2005
DEFINITION AGENCOURT 40973844 NIH MGC 281 Homo sapiens cDNA clone
 IMAGE:7782009 3', mRNA sequence.
ACCESSION CX757659
VERSION CX757659.1 GI:58054315
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 796)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Meri Firpo
 cDNA Library Preparation: Express Genomics

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 <http://image.llnl.gov>
 Plate: LHAM5939 row: p column: 07
 High quality sequence start: 16
 High quality sequence stop: 792.
 Location/Qualifiers
 1..796

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:7782009"
/tissue_type="pluripotent cell line derived from
blastocyst inner cell mass"
/lab_host="DH10B Tona"
/clone_lib="NIH MGC 281"
/note="Organ: Blastocyst; Vector: pExpress-1; Site 1:
EcoRV; Site 2: NotI; RNA obtained from pluripotent cell
line derived from blastocyst inner cell mass (cell line
HSP-6, NIH Registry designation UC06. Positive for OCT4
expression by RT-PCR, positive for SSEA-3, SSEA-4,
Tra-1-81, Tra-1-60 by immunofluorescence. Negative for
SSEA-1 by immunofluorescence Passage 62. cDNA was primed
using oligo-dT primer.
5'-pGACTGATCTGATGCGAGCGCGCCGCC(1)25-3' and cloned into
the EcoRV/NotI sites of pExpress-1. Size-selection >1.25
kb resulted in an average insert size of 2.0 kb. This
primary library is normalized (non-normalized primary
library is NIH MGC 280) and was constructed by Express
Genome (Frederick, MD). Note: this is a Mammalian Gene
Collection Library."

ORIGIN

Query Match 14.5%; Score 167.6; DB 8; Length 796;

Best Local Similarity 54.4%; Pred. No. 1.8e-41;

Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

421 CCTAAGATTGTTGTCAGAGCATCGTGTATTTATTTGTCCTTATAGATCGTGA 480
122 CCGGAGGTTGAGCCCGCTCCGAGACGATCATTTGCTCATGTCGCCGGAG 181
481 GCACATTGAGATATATGCTCCCAATTGTCATCTGCTGCCCAACAATTTGAC 540
182 CACCACTGCGCTGCTGCTTACCACTGCCCTTTCAGAGCCAGACGCTGCT 241
541 TATGCAATTTTCAATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
242 TATGCAATTTTCAATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 301
601 AACGTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
302 AACGTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361
661 GTGCAATTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 717
362 GTGCACTTCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 421
718 CATGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 777
422 CATGAGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
778 GGAATCAGTGAATCAAAAGATCACTGAAGAAATCAATGATTTTGAATGAT 837
482 GGAATCAGTGAATCAAAAGATCACTGAAGAAATCAATGATTTTGAATGAT 541
838 TGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 897
542 TGGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 601
898 GTTTCAAGATTCGAGCAAAATGAGATTAATAATGATTAAGATTCGAGCAAG 957
602 ATCTCTCGGCCCCCAGATCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
958 ACGAATCCAGTTAATTAATGCGCTAACAATTAATGCGGAGGAGGAGGAGGAG 1017
662 GCGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 721
1018 CGTGAAGGCTTGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1077
722 CAAAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 781

QY 1078 AC 1079
DB 782 AC 783

RESULT 7
CD514077
LOCUS
DEFINITION
AGENCOURT 14373501 NIH MGC 179 Homo sapiens cDNA clone
IMAGE:30354133 5', mRNA sequence.
ACCESSION
CD514077
VERSION
CD514077.1 GI:31445795
KEYWORDS
EST.
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 879)

REFERENCE
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10X7 Bethesda, MD 20892
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: NDAM472 row: c column: 06
High quality sequence sloop: 659.
Location/Qualifiers

FEATURES

source

1..879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30354133"
/tissue_type="pituitary"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH MGC 179"
/note="Organ: Brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV
(destroyed); Site 2: NotI; Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.1 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 14.5%; Score 167.6; DB 6; Length 879;

Best Local Similarity 54.4%; Pred. No. 1.9e-41;

Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

421 CCTAAGATTGTTGTCAGAGCATCGTGTATTTATTTGTCCTTATAGATGTTGA 480
45 CCGGAGGTTGAGCCCGCTCCGAGACGATCATTTGCTCATGTCGCCGGAG 104
481 GCACATTGAGATATATGCTCCCAATTGTCATCTGCTGCCCAACAATTTGAC 540
105 CACCACTGCGCTGCTGCTTACCACTGCCCTTTCAGAGCCAGACGCTGCT 164
541 TATGCAATTTTCAATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
165 TATGCAATTTTCAATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 224
601 AACGTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
225 AACGTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 784

OY	661	GTGCATTTACTGCCCGAAGATGACCCGTAACTGTACAGTGT---CCAAATTCACACAGT	717
Db	285	GTGAACCTCTTGGCCAGAAATGACCACAACTGTATGTGTGAGACCCCGGGAGCCCGC	344
OY	718	CATATGAGTGTAGCGATCGATTAATTCATTAATACTTCCATATTCGGCCATCTTGGGC	777
Db	345	CATGTGGCCGTGTGCATGAACAAGTTTGGATACAGCCCTCCGTTACCCCAATCACTTGGGA	404
OY	778	GGAATCACTGCATTAACAAAGATCACTGAAAGAAATCAATGGATTTTGCATATGATTTT	837
Db	405	GGAGTTCAGAGCACTTACTCTGACCCAGTACCTGAAAGAAATGGTTTCCCAATGAATAAC	464
OY	838	TGGGGTTGGGGCGGAGAGACGACGATTTGGCGACGACAAATCGATGGCTGCACTGAAA	897
Db	465	TGGGCTTGGGTTGGTGAAGATGACGACATTGCTACAGGGTGGCGCTTGGCTGGATGAAG	524
OY	898	GTTCACAGTATCCGACACAAATTTGCACGATATAATAATGATTAAGCACTCGACGGACGG	957
Db	525	ATCTCTCGGCCCCCACCACATCTGTAGGACACTATATGATGTGTAAGCACCGAGAGATTAAG	584
OY	958	ACGAATCCAGTTAATAATAATGCCCCGTACAAATAATATGGGCGCAAGACAGCCGATGCA	1017
Db	585	GGCAATGAGGAAAAATCCCCACAGATTTGACCTCTGTGTCGGTACCCCAAAATTCCTGAGCG	644
OY	1018	CGTACGGCGCTPAAGCAATCTGAAATTAAGCTCGTAATCTGGAATTAAGACCTCTAC	1077
Db	645	CAAGATGGGATGAACTCACTGACATACCAAGTTGCTGGCTCCAGAGAGCTGGGCGCTTTAT	704
OY	1078	AC 1079	
Db	705	AC 706	

LOCUS	BO957788	955 bp	mRNA	linear	EST 21-AUG-2002
DEFINITION	AGENCOURT.8772708 NIH_MGC_18 Homo sapiens cDNA clone IMAGE:5372348				
ACCESSION	5'', mRNA sequence.				
VERSION	BO957788				
KEYWORDS	BO957788.1	GI:22373266			
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 955)				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-r@mail.nih.gov Tissue Procurement: DCTD/DTF/Gazdar cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: http://image.llnl.gov Plate: L1CM2547 row: p column: 13 High quality sequence stop: 617.				

FEATURES

```

1:organism="Homo sapiens"
2:organelle="mitochondrion"
3:mol_type="mRNA"
4:db_xref="taxon:9606"
5:clone="IMAGE:6372348"
6:tissue_type="large cell carcinoma"
7:lab_host="RDH10B (phage-resistant)"
8:clone_lib="NHG_MGC_18"
9:note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2:
10: EcoRI; cDNA made by oligo-dT priming. Directionally cloned
11: into EcoRI/XhoI sites using the following 5' adaptor:

```

GGCAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

ORIGIN	
Query Match	14.5%; Score 167.6; DB 5; Length 955;

Query Match	14.5%	Score 167.6	DB 5	Length 955
Best Local Similarity	54.4%	Pred. No. 1.9e-41		
Matches 360; Conservative	0	Mismatches 299;	Indels 3	Gaps 1

OY	421	CTTAAGGATGTTGTTGCAAGGACATCGGTTGGCTATTTATTTGGCCCTTAATGAGATCGTGA	480
Db	79	CTCGAGGTTGTGAGCCCGCTCCGAAACAGCCATCATTTGCTCTCATGTGCCCCGGAG	138
OY	481	GCACATTTGGAATATATGCTCCACAATTTGCACTGTTGCTCGCAAAACAATTTGAC	540
Db	139	CACCACCGGCGCTGCTGCTCTACCACTGCAACCCCTTTCTTGACGCCACAGCACTTGCT	198
OY	541	TATGCAATTTTCATTTGTGAGCAAGTGGCAATCAGACGTTTAAATCGCGGAAATCTATG	600
Db	199	TATGGCATCTATATGTCATCCACAGGCTGGAATGAAACATTTAACAGGCAAACTGTTG	258
OY	601	AACGTTGGATACAGAGTATGATCAAGGCTCAACCCATGAGGAGTGCTTCATCTTTCATGAT	660
Db	259	AACGTTGGGCGGCGAGAGGCGCCCTCGATGTAAGAAGTGGAGCTGCGTTCTTGCAGAT	318
OY	661	GTGCAATTTACTGCCCGAAGATGACCGTAACTGTATACAGCTGT---CCAATTCAAACACGT	717
Db	319	GTGAGACCTCTTGGCAGAAATGACCAACATCTGTATGTGTGTGACCCCGGGGACCCCGC	378
OY	718	CATATGAGTGTAGCGATGATTAATTCATATATATACTTCATATTTGGGCATCTTGGC	777
Db	379	CATGTTGCGGTTGCTATGAAACAAGTTTGGATATACAGCTCCCGTACCCCGATATTGGGA	438
OY	778	GGAATCAAGTGCATPAACAAAAGATCACCTGAAGAAATCAATGGAATTTGCAATGATTTT	837
Db	439	GGAGTCTCAGCACTTACTCTGACCAAGTACTGGAAGATGAAATGCTTCCCAATGAATATAC	498
OY	838	TGGGTTTGGGGCGGAGAGGACGACGATTTGGCGACGAGAACATGATGGCTGTGATGAA	897
Db	499	TGGGCTGTGGGTGGTGGAGATGACGACATTTGTCACAGGTCGCGCTGTGGCTGGATGAG	558
OY	898	GTTCACAGATATCCGACACAATTTGCAAGATATATAAATGATTAAAGCACTCGACGGAAAG	957
Db	559	ATCTCTGGGCCCCCACAATCTGTAGGACATATATAAGATGTGAAGACCCGAGAGATTAAG	618
OY	958	ACGAATCCAGTTAATTAATGCGGCTACAAATAATGCGGCCAAAGAGCGCCGATGACA	1017
Db	619	GGCAATGAGGAAATCCCCACAGATTTGACCTCCGTGTCCTTACCACAAATTCCTGGACG	678
OY	1018	CGTACGCGCCTTAAGCAATCTGAAGTATTAAGCTGTAATCTGGAATTTGAGCCTCTCTAC	1077
Db	679	CAAGATGGGATGAATCACTGACATTAACAAGTTGTGCTGCGTCGAGAGGCTGGGCGCTTTTAT	738
OY	1078	AC 1079	
Db	739	AC 740	

RESULT 5

RESULT 9

DQ035706
LOCUS

DQ035706				
LOCUS	DQ035706	1196 bp	DNA	linear
				GSS 02-JUN-2005

DEFINITION

DEFINITION Homo sapiens BAGALT3 gene, VIRUTAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION
VERSION

ACCESSION	DQ035706
VERSION	DQ035706.1
	GI:66886915

KEYWORDS
SOURCE

KEYWORDS	GSS.
SOURCE	Homo sapiens (human)

ORGANIZATIONAL

ORGANISM	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBLISHED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
gene
ORIGIN

1 (bases 1 to 1196)
Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Pieidel-Alon, A., Tenebaum, D.M., Civeello, D., White, T.J., Smitsky, J.J., Adams, M.D. and Cargill, M.
A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
(er) PLoS Biol. 3 (6), E170 (2005)
15869325
2 (bases 1 to 1196)
Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Pieidel-Alon, A., Tenebaum, D.M., Civeello, D., White, T.J., Smitsky, J.J., Adams, M.D. and Cargill, M.
Direct Submission
Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.

Location/Qualifiers
1. 1196
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
<1..>1196
/gene="BAGALT3"
/locus_tag="HC4901"

Query Match 14.5%; Score 167.6; DB 11; Length 1196;
Best Local Similarity 54.4%; Pred. No. 2.1e-41;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

421 CCTAAGATTGTTGTCAGAGCATGCTGTTGCTATTATTTGCTTATAGATGCTGAA 480
360 CCGCAGGTTGTGAGCCCGCTCCCAAGACGCATATGTGCTATGCTGCCGGAG 419
481 GCACATTTGAGATATGCTCCACATTTGCACTGCTGCTGCCAACAACATTTGAC 540
420 CACCACTGCGCTGCTGCTTACACACTGCACTCTTCTTGAGGGCCAGCAGCTTGCT 479
541 TATGCAATTTTCATTGTGAGCAAGTGGCGAATCAGACGTTTATGCGGAAACTAATG 600
480 TATGCACTTATGTATCATCACACGCTGGAATATGAACATTTAAGGGCAAACTGTTG 539
601 AACGTTGATAGACGTAGCATGACGCTCTTACCCATGCGAGTCTTCACTTTTCAT 660
540 AACGTTGGGGTGGAGAGGCCCTGCGTGAATGAAGTGGAGCTGCTGTTTGGACGAT 599
661 GTGCAATTTACTGCGCCGAGATGACCGTAACTGTACACGTT--CCAAATTCACACGAT 717
600 GTGCACTCTTGGCAGAAATATGACCAATCTGTATGTGTGATACCCCGGGGACCCCGC 659
718 CATATGATGTAGCGATCATTAATTCATTTAATCACTTCCATATTCGCGATCTTTCGC 777
660 CATGTGCGGTGTCTATGAACAAGTTTGGATACAGCTCCCGTACCCCAAGTACTTCGGA 719
778 GGAATAGTGCATTAACAAAGTACCTGAGAAATCAATGATTTTTCGATGATTTT 837
720 GGAATAGTGCATTAACAAAGTACCTGAGAAATCAATGATTTTTCGATGATTTT 779
838 TGGGGTGGGGGAGAGGACGATTTGGCGAGAGAAACATCATGCTGGGACTGGA 897
780 TGGGGTGGGGTGGTGAAGATGACATTTGCTACACAGGGTGGCTGGTGGGATGAAG 839
898 GTTTCAGATATTCGACACAAATTCGACGATATTAATGATTAAGCACTGACGAGAGCG 957
840 ATCTCTGGGCCCCCAGATCTGTAGACACTATTAATGATGGTGAAGACGAGAGATAG 899
958 AGCAATTCAGTTTATTAATGCGCTACAAATTAATGGGCCAAAGAGGCGCATGGA 1017
900 GGCAATGAGAAATATCCACAGATTTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 959

QY 1018 CCGTACGGCCTTAAGCAATCTGATATTAAGTGAATGAGCCTCTTAC 1077
DB 960 CAAATGGGATGAATCACTGACATACCAAGTTGCTGGCTCGAGAGCTGGGCTCTTAT 1019
QY 1078 AC 1079
DB 1020 AC 1021

RESULT 10
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CR614082
full-length cDNA clone CS0D005YK01 of T cells (Jurkat cell line)
Cot 10-normalized of Homo sapiens (human).
CR614082
HTC; CNSLT_cDNA.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1590)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ invitrogen Corporation 1600 Faraday Avenue
2 (bases 1 to 1590)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
Location/Qualifiers
1. 1590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D005YK01"
/tissue_type="T cells (Jurkat cell line) Cot 10-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 14.5%; Score 167.6; DB 4; Length 1590;
Best Local Similarity 54.4%; Pred. No. 2.3e-41;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;

421 CCTAAGATTGTTGTCAGAGCATGCTGTTGCTATTATTTGSCCTTATAGATGCTGAA 480
282 CCGCAGGTTGTGAGCCCGCTCCGAAACGCATATGTGCTCATGTGCTGCCGGAG 341
481 GCACATTTGAGATATGCTCCACATTTGCACTGCTGCTGCCAACAACATTTGAC 540
420 CACCACTGCGCTGCTGCTTACACACTGCACTCTTCTTGAGGGCCAGCAGCTTGCT 401
541 TATGCAATTTTCATTGTGAGCAAGTGGCGAATCAGACGTTTATGCGGAAACTAATG 600
402 TATGCACTTATGTATCATCACACGCTGGAATATGAACATTTAAGGGCAAAACTGTTG 461
601 AACGTTGATAGACGTAGCATGACGCTCTTACCCATGCGAGTGGTCACTTTTATGAT 660
462 AACGTTGGGGTGGAGAGGCCCTGCGTGAATGAAGTGGAGCTGCTGTTTGGACGAT 521
661 GTGCAATTTACTGCGCCGAGATGACCGTAACTGTACACGTT--CCAAATTCACACGCT 717

Db	522	GTGACCTCTTCCAGAAAAAGACCACAACTGTATGTGTGTGACCCCCGGGACCCCGC	581
Qy	718	CATATGAGTGTAGCGATCGATTAATTCAATTATATACTTCATATGTGGGATCTTGCGC	777
Db	582	CATGTTCGCCGTGTCTATGAACAAGTTTGGATACAGCTCCCGTACCCCACTACCTGGGA	641
Qy	778	GGATTCAGTGCATCAACAAAGATCACCTGAAAGAAATCATGGAATTTTGGATGATTTT	837
Db	642	GGAGCTTCAGACTACTCTCTGACACAGTACCTGAAAGATGAATGGCTTCCCAATGAATAC	701
Qy	838	TGGGCTTGGGGCGGAGAGACGACGATTTTGGCGACGAAACATCGATGGCTGACTGAAA	897
Db	702	TGGGGCTGGGGTGGTGGAGGATGACGACATTCCTACCAAGGTGGCGCTGGCTGGATGAAG	761
Qy	898	GTTTCAGATATTCGACACAAATTGCAAGATATATAATGATTAAGCACTGACGGAACG	957
Db	762	ATCTCTCGGCCCCCACCATCTGTGTAGGACACTATATAGATGTGAAGCACCCGAGGATPAG	821
Qy	958	ACGATTCAGTTATTAATATGCGCGCTACAAATTAATGGGCGCAACGGAAGCGCGATGACA	1017
Db	822	GGCAATGAGAAAAATCCCCACAGATTTTGACTCTCGTGTCCGTATCCAGAAATTCCTGGACG	881
Qy	1018	CGTACGCGCCTPAGCAATCTGAGATATTAAGCTCGTAAATTCGGAATTTGAACCTCTCTAC	1077
Db	882	CAAGTGTGGATGAATCACTGACATACACAGTTGTGCTCGAGAGACTGGGCGCTCTTAT	941
Qy	1078	AC 1079	
Db	942	AC 943	

RESULT	11
CRC612341	
LOCUS	1860 bp mRNA linear HTC 21-JUL-2004
DEFINITION	Full-length cDNA clone CSDDK011YM18 of Hela cells Cot 25-normalized
ACCESSION	CRC612341
VERSION	CRC612341.1 GI:50493148
KEYWORDS	HTC; CNSLT cDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
TITLE	Homidae; Homo.
JOURNAL	1 (bases 1 to 1860)
REMARK	Li, W.B., Gruber, C., Jessee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished
	Contact : Peng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1860)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (20-JUN-2004) Genoscope - Centre National de Sequenace :
JOURNAL	BP 191 91006 EVRY cedex - PRANCE (E-mail : sequefgenoscope.cns.fr
COMMENT	- Web : www.genoscope.cns.fr) 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	Location/Qualifiers
source	1..1860
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/clone="CSDDK011YM18"
	/tissue_type="Hela cells Cot 25-normalized"
	/plasmid="pCMVSPORT_6"
ORIGIN	
Query Match	14.5%; Score 167.6; DB 4; Length 1860;

Best Local Similarity 54.4%; Pred No. 2.4e-41; Matches 360; Conservative 0; Mismatches 229; Indels 3; Gaps 1;									
QY	421	CCTAAGGATTTGTTGTCAGAGGACATCGTGTGCTATTATTTGTGCCCTTATAGAGATCGTGA	480						
Db	562	CTGTCAGATGTTGAGGCCCCCTCCCGAAGCCATCATTTGTGCTCATGTGCCCCGGAG	621						
QY	481	GCACATTGGAATTAATGCTCCCAAAATTTCGACGTGTGCTCGCCAAACAATTTGAC	540						
Db	622	CACCACTGCGCCTGCTGCTCTTACCACTGACCCCTCTTCGAGCGCCAGAGCTTGCT	681						
QY	541	TATGCAATTTTTCATTGTGAGCAAGTGGCAATCAGCGTTTAAATCGCGGAAACTAATG	600						
Db	682	TATGCACTATATGTCATCCACAGGCTGTGAATGGAACATTTAAACAGGGCAAACTGTG	741						
QY	601	AACCTTGGATACGACGTAGCATCAGCGCTCTACCCAGGAGGCTTCATCTTTCATGAT	660						
Db	742	AACCTTGGGCTGCGAGGCGCCTGCGATGAGAGTGGAGCTGCTGTCTTTCACGAT	801						
QY	661	GTGCATTTTACTGCCCGAAGATGACCGTAACTGTACACGTGT--CCAAATTCAACAACGT	717						
Db	802	GTGACCTCTTTCGCAAAATGACACACATCTGTATGTGTGTACCCCGGGAGCCCGC	861						
QY	718	CATATGAGTGTAGCGATCGATTAATTCATTATTAACCTTCATATTTGGCGATTTGGC	777						
Db	862	CATTTGTCGGTGTCTATGAAACAAGTTTGTGATACAGCCTCCGTAACCCCACTAGCTTGGGA	921						
QY	778	GGATTCAGTGCACATCAAAAGATCACCTGAGAAATTCATGGATTTTTCGAATGATTTT	837						
Db	922	GGAGTCTCAGCACTTACTCTGACCACTAGTGAATGATGAGCTTCCCAATGAATAC	981						
QY	838	TGGGGTTTGGGCGGAGAGGACGACGATTTGGCGACGAAACATGATGCTGACTGA	897						
Db	982	TGGGGCTGGGGTGTGTGAGATGACGACATTTGCTACAGGGGTCGCTGTGGCTGATGAAG	1041						
QY	898	GTTTCAAGATATTCGACACAAATTCGACGATATTAATATGATTAGACTCGACGGAACG	957						
Db	1042	ATCTCTGGGCCCCCACATCTGTAGGACACTATTAAGATGTGGAAGCACCGAGAGATPAG	1101						
QY	958	ACGAATCAAGTTATTAATGCGGCTACAAAATAATGCGCCAAAGAGCGCCGATGACA	1017						
Db	1102	GGCAATAGGAAATTCGCCACAGATTGACCTCCTGTGCTCGTACCGAAATTCCTGAGC	1161						
QY	1018	CGTAGCGGCTTACGCAATCTGGAATATPAGCTCGTAAATCTGGAATTGAACCTCTCTAC	1077						
Db	1162	CAGATGGGATGAACCTCACTGACATACCAAGTTGCTGCTCGAGAGCTGGGGCTCTTTAT	1221						
QY	1078	AC 1079							
Db	1222	AC 1223							

RESULT 12				
LOCUS	CR607487	1870 bp	mRNA	linear
DEFINITION	Full-length cDNA clone CSDDC015Y21 of Neuroblastoma Ccr			
	25-normalized of Homo sapiens (human).			
ACCESSION	CR607487			
VERSION	CR607487.1	GI:50488294		
KEYWORDS	HTC; CNSLT cDNA.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1870)			
AUTHORS	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.			
TITLE	Full-length cDNA libraries and normalization			
JOURNAL	Unpublished			
REMARK	Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue			
REFERENCE	2 (bases 1 to 1870)			

AUTHORS

Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segr@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

Location/Qualifiers
1..1870
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D015YB21"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 14.5%; Score 167.6; DB 4; Length 1870;
Best Local Similarity 54.4%; Pred. No. 2.4e-41;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;
OY 421 CCTAAGATTGTGTGTCAGAGCATCGTGTCTATTATTGTGCGCTATAGAGTCGTGAA 480
DB 574 CCGGAGGTTGTGAGCCCGCTCCGACAGCCATCATTTGTGCTCATGTGCGCGGAG 633
OY 481 GCACATTTGAAATATGCTCCCAATTTGCACTGCTGCTGCCAAACAATTTGAC 540
DB 634 CACCACTGCGGCTGCTGCTTACCACTGCACTCCCTTTTGAGGCGCAGCAGCTGTG 693
OY 541 TATGCAATTTTCAATTTGTGAGCAAGTGGCAATCAGACGTTTAATGCGGGAACATG 600
DB 694 TATGCAATTTTCAATTTGTGAGCAAGTGGCAATCAGACGTTTAATGCGGGAACATG 753
OY 601 AACGTTGATACGACGTAGCATCAGCTCTTACCCATGCGACGTCTTCACTTTTCATG 660
DB 754 AACGTTGATGAGGAGGAGGCGCTGCTGATGATGATGATGATGATGATGATGATGAT 813
OY 661 GTGCAATTTACTGCGGCAAGATGACCTGATACAGTGT---CCAATTCAACCAAGT 717
DB 814 GTGCACTCTTTCGCAAAATGACCAATCTGTATGTGTGATGATGATGATGATGATGAT 873
OY 718 CATGAGAGTGAAGCATGATTAATCAATTAATCAATTAATGCGGATCTTCCGAC 777
DB 874 CATGTTGCCGTTGCTATGAAACAAGTTGATACAGCTCCCGTACCCCAAGTACTTCGA 933
OY 778 GGAATCAGTCACTAACAAGATCACTGAAGAAATCAATGATTTTCAATGATTTT 837
DB 934 GGAATCTGAGCACTTACTCTGACGATACCTGAATGATGATGATGATGATGATGATGAT 993
OY 838 TGGGTTGGGGCGAGAGACGACGATTTGGCGACGAAACATGATGCTGACCTGA 897
DB 994 TGGGTTGGGGCGAGAGACGACGATTTGGCGACGAAACATGATGCTGACCTGA 1053
OY 898 GTTTCAAGATTCGACGAAATTTGACGATTAATTAATTAATTAATTAATTAATTAATTA 957
DB 1054 ATCTCTCGGCCCCCATCTGTGAGGACATTAATTAATTAATTAATTAATTAATTAAT 1113
OY 958 ACGAATCCAGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1017
DB 1114 GGCATAGAGAAATATCCACAGATTTGACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1173
OY 1018 CGTACGCTTAAAGCAATTTGAATTAAGTCTGTAATTTGGAATTTGAAGCTCTCTAC 1077
DB 1174 CAAAGTGGATGAATCACTGACATACAGATGCTGCTGAGAGCTGGGGGCTCTTTAT 1233
OY 1078 AC 1079
DB 1234 AC 1235

RESULT 13

CR592876

LOCUS

DEFINITION

of Homo sapiens (human).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

Unpublished

Contact : Feng Liang Email : fliang@life.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue

2 (bases 1 to 1876)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

Location/Qualifiers

1..1876

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1008Y21"

/tissue_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 14.5%; Score 167.6; DB 4; Length 1876;
Best Local Similarity 54.4%; Pred. No. 2.4e-41;
Matches 360; Conservative 0; Mismatches 299; Indels 3; Gaps 1;
OY 421 CCTAAGATTGTGTGTCAGAGCATCGTGTCTATTATTGTGCGCTATAGAGTCGTGAA 480
DB 575 CCGGAGGTTGTGAGCCCGCTCCGACAGCCATCATTTGTGCTCATGTGCGCGGAG 634
OY 481 GCACATTTGAAATATGCTCCCAATTTGCACTGCTGCTGCCAAACAATTTGAC 540
DB 635 CACCACTGCGGCTGCTGCTTACCACTGCACTCCCTTTTACAGGCGCAGAGCTGTG 694
OY 541 TATGCAATTTTCAATTTGTGAGCAAGTGGCAATCAGACGTTTAATGCGGGAACATG 600
DB 695 TATGCAATTTTCAATTTGTGAGCAAGTGGCAATCAGACGTTTAATGCGGGAACATG 754
OY 601 AACGTTGATACGACGTAGCATCAGCTCTTACCCATGCGACGTCTTCACTTTTCATG 660
DB 755 AACGTTGATGAGGAGGAGGCGCTGCTGATGATGATGATGATGATGATGATGATGAT 814
OY 661 GTGCAATTTACTGCGGCAAGATGACGTTAACGTTACAGTGT---CCAATTCAACCAAGT 717
DB 815 GTGCACTCTTTCGCAAAATGACCAATCTGTATGTGTGATGATGATGATGATGATGAT 874
OY 718 CATGAGAGTGAAGCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 777
DB 875 CATGTTGCCGTTGCTATGAAACAAGTTGATACAGCTCCCGTACCCCAAGTACTTCGA 934
OY 778 GGAATCAGTCACTAACAAGATCACTGAAGAAATCAATGATTTTCAATGATTTT 837
DB 935 GGAATCTGAGCACTTACTCTGACGATACCTGAAGATGAATGAGCTTCCCAATGAATAC 994

[illegible]

RESULT 14	
CR599499	
LOCUS	1901 bp mRNA linear HTC 21-JUL-2004
DEFINITION	full-length cDNA clone CS0DD008Yp16 of Neuroblastoma Cot
ACCESSION	CR599499
VERSION	CR599499.1 GI:50480306
KEYWORDS	HTC; CNS12 cDNA.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
	Hominidae; Homo.
	1 (bases 1 to 1901)
REFERENCE	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
AUTHORS	Full-length cDNA libraries and normalization
TITLE	unpublished
JOURNAL	Contact : Feng Liang Email : liang@lifetech.com URL :
REMARK	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

REFERENCE	2 (bases 1 to 1901)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (20-Jul-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES	Location/Qualifiers
source	1. 1901 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CSDD0081P16" /tissue_type="Neutroblastoma Cot 50-normalized" /plasmid="pCMVSPORT_6"
ORIGIN	

Query Match	14.5%	Score 167.6;	DB 4;	Length 1901;
Best Local Similarity	54.4%	Pred. No. 2.4e-41;		
Matches 360; Conservative	0;	Mismatches 299;	Indels 3;	Gaps 1

QY 421 CCTAAGGATTTGGTTCGCAAGGACATCGTGTGCATTAATTGTGCCCAATGAGATCGTGA 480

Db 583 CCTGACAGTTGTGAGGCCCGCCGTCGCCAAGACATATTGTGCTCATCGTGC CGGGAG 642

QY 481 GCACATTTGAGATTAATGCTCCACAATTTGCACTGTGTGCTGCCAACAACAATTTGAC 540

Db 643 CACCACTTGCGCTGCTGCTCTACACCTGCACCCCTTCTTGACGCGCCAGGACGCTTGT 702

OY	541	TATGCATTTTTCATTGTGGAGCAAGTGGCGAATTCAGAGTTTAAATGCGGGAAACTATG	600
Db	703	TATGGCACTATATGTCATTCACACAGCTGGAGATGAACATTTTAAACAGGCAAACTGTG	762
OY	601	AACCTGGATTAACGACGTAGCATCAAGCCTCTACCCATGGCAGTGCCTTCATCTTTCATGAT	660
Db	763	AACCTTGGGGTGGCAGAGAGCCCTCGATGAAAGTGGGACTCTGTTTCTTGACGAT	822
OY	661	GTGCATTTTACTGCCCCGAAGATGACCGTAACTGTACACGTGT---CCAATTCAACCAAGT	717
Db	823	GTGACACTCTTGGCCAGAAATGACCAACATCTGATGTGTGTGACCCCGGGAGCCCGC	882
OY	718	CATATGATGTAGGAGATCGATTAATTTCAATATATTAATCTTCATATTTGGAGATCTTGGC	777
Db	883	CATGTGCGGTGCTTATGACAAAGTTTGGATACAGCTCCCGTACCCCACTACTTGGGA	942
OY	778	GGAATCAGTGCACTAACAAAGATCACCTGAAGAAAAACAATGGATTTTTCGATGATTTT	837
Db	943	GGAGTCTCAGACCTTACTCTGACCAAGTACTGAAGATGATGGCTTCCCAATGAATAC	1002
OY	838	TGGGCTTGGGGCGGAGAGACGACGATTTTGGCGACGAGACATGATGCTGGACTGAA	897
Db	1003	TGGGGCTGGGGTGGTGGAGATGACGACATTTGCTCACAGGGGTGCGCTGGCTGGGATGAAG	1062
OY	898	GTTTTCAGATATTCGACACAAATTTGACGATATTAATTAATGATTAAGCACTGACGGAAAGC	957
Db	1063	ATCTCTGGCCCCCACCACATCTGTGAGCACTATTAAGATGGTGAAGCACCGAGAGATTAAG	1122
OY	958	ACGATTCAGTAAATTAATGACCGCTACAAATAATGGCCAAAGAAAGCCCGCATGACA	1017
Db	1123	GGCAATAGGAAAAATCCCAACAGATTTGACCTCGTGGTCCGTACCCAGAAATTCCTGGAGC	1182
OY	1018	CGTAGCGCCTTAAGCATCTGAAGTATTAAGTCTGTAATCTTGGAAATTGAACCTCTCTAC	1077
Db	1183	CAAGTATGGATGAACACTCACTGACATACCAAGTATGCTGAGCTCGAGAGCTGGGGCTCTTAT	1242
OY	1078	AC 1079	
Db	1243	AC 1244	

RESULT	15
AJ740626/c	
LOCUS	AJ740626 673 bp mRNA linear EST 17-FEB-2005
DEFINITION	AJ740626 riken1 Gallus gallus cDNA clone 19hs87, mRNA sequence.
ACCESSION	AJ740626
VERSION	AJ740626.1 GI:53906003
KEYWORDS	EST
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus;
REFERENCE	1 (bases 1 to 673).
AUTHORS	Caldwell,R.B., Kierzek,A.M., Arakawa,H., Bezubov,Y., Zaim,J., Fiedler,P., Kutser,S., Blagoderatski,A., Kostova,D., Koter,M., Plichy,J.J., Caminci,P., Hayashizaki,Y. and Buerstedde,J.M. Full-length cDNAs from chicken bursal lymphocytes to facilitate gene function analysis <i>Genome Biol.</i> 6 (1), R6 (2005) 15642098
JOURNAL	Contact: Caldwell RB
PUBMED	GSF - Forschungszentrum, Institut fuer Molekulare Strahlenbiologie Ingolstaedter landstr. 1, D-85764 Neuherberg, GERMANY.
COMMENT	Location/Owning laboratory
FEATURES	

FEATURES	Location/Qualifiers
source	1..673
	/organism="Gallus gallus"
	/mol_type="mrna"
	/db_xref="taxon:9031"
	/clone="19h387"
	/cell_type="bursal lymphocyte"
	/dev_stage="2-3 weeks old"

ORIGIN /clone lib="riken1"
/note="CB inbred strain"

Query Match 14.3%; Score 165; DB 1; Length 673;
Best Local Similarity 56.6%; Pred. No. 1.2e-40;
Matches 306; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

```
QY 393 GGACACGACGCGCGGTGGACATGATGCTTAAGATTGTGTGCAAGGATCGGTG 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 542 GGACATTAAACGAGAGACACTGGAAACCCAAAGCTGTAGCCACGATGAGAGTGGC 483
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 453 TATTATTGCGCCCTATGAGATCGTGAAGACATTTGAGATATATGCTCCACATTTGCA 512
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 482 GATCATCATCTTCTTTTCGTAATCGTCAGAGCATCTTCCAAATTTCTTCGGGCACTGTAT 423
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 513 CTGTTGCTCGCCAAACAACTGATGACTATGCAATTTTCATTGTGAGCAAGTGGCGAA 572
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 ACCGATGTTGCAGAAAGCAGCGGCTGGAAATTTGCTTCTATGTTGTTGAACAGACAGGTAC 363
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 573 TCAGACGTTTATCCGCGGAAACTAATGAAGTTGATACGACGTACATCAAGCTCTTA 632
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 ACAACCTTTTAACTGTCATGCTTTTAACTGTGCTTCAAGAGGCCATGAGAGATGT 303
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 633 CCCATGGCAGTGTCTCATCTTCATGATGTGATTTTACTGCCGAAGATGACCGTAACT 692
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 TGCTGGGACCTGCAATATATTTTCATGATGTGATCACTTACCTGAATAATGACCGAAATTA 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 693 GTACAGTGTCCAAATTCACACGTCATATGAGTGTAGCGATCGATTAATTTATATA 752
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 242 TTACGAGATGTGAGAAATGCGACGTCATTTTGCAGCAAAAGTTGACAAATACATGTACAT 183
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 753 ACTTCCATATTTGGGGGATTTTGGGGGATTCAGTGCATTAACAAAAGATCAGCTGAAGAA 812
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 182 TCTTCCATACAAATGAGTCTTGTGTGTGTAGTGTGACTGACAGTGGAACTTCAAGAA 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 813 AATCAATGGAATTTTGAATGATTTTGTGGGTTGGGGCGGAGAGACGACGATTTGGCGAC 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 122 GATTAAATGGAATTTCCAAATGCTTCTGSGGTTGGGGTGGAGAAAGATGATGATCTTTGGAA 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 873 GAGAAATGATGCTGGAAGTGAAGTTTCAAGATATCCGACACAATTTGACGATATATA 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 CAGGGTTCACTATGCTGTGATACACGTAAACAGACCAAGAGGAGACTTGGGAAATATACAA 3
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 933 A 933
    |
DB 2 A 2
```

Search completed: January 8, 2006, 22:26:44
JDB time : 5197 secs

THIS PAGE BLANK (USPTO)